

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 49.0254 Seconds
(without alignments)
4211.715 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMYRR.....QHKRVSQFLDDPSTAEVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5531.5	49.6	2201	2 A54774	ATP binding cassette
2	2863.5	25.7	1529	2 A59189	ATP-binding cassette
3	2694.5	24.2	1472	2 B54774	ATP binding cassette
4	2503	22.5	1704	2 S71363	probable ATP-binding
5	2502	22.5	1704	2 A59188	ATP-binding cassette
6	1829	16.4	1802	2 T33783	hypothetical prote
7	1813	16.3	1816	2 A84845	probable ABC trans
8	1603	14.4	1317	2 C88925	protein F33E11.4 [
9	1600.5	14.4	1447	2 T15200	hypothetical prote
10	1408.5	12.6	1758	2 F85559	protein C48B4.4b [
11	1407	12.6	1767	2 S60124	transport protein
12	1405	12.6	1704	2 T42749	ATP-binding cassette
13	1376	12.3	1246	2 T00826	hypothetical prote
14	1345	12.1	269	2 T46467	hypothetical prote
15	1038.5	9.3	1564	2 T27121	hypothetical prote
16	937.5	8.4	1431	2 T22748	hypothetical prote
17	901.5	8.1	373	2 T47150	hypothetical prote
18	741.5	6.7	1011	2 T07712	probable ABC-type
19	665	6.0	895	2 T07714	probable ABC-type
20	650	5.8	900	2 T07717	probable ABC-type
21	625.5	5.6	722	2 T07716	probable ABC-type
22	618.5	5.6	925	2 T07713	probable ABC-type
23	539	4.8	1336	2 T18288	ABC transport prot
24	435	3.9	582	2 H95950	probable ABC trans
25	429	3.8	330	2 S27707	daunorubicin resis
26	423.5	3.8	664	2 T07715	probable ABC-type
27	419.5	3.8	342	2 T36505	probable ABC-type
28	410	3.7	339	2 AE2102	ABC transporter AT
29	408.5	3.7	333	2 D72492	probable ABC trans

RESULT 1

A54774
ATP binding cassette transporter ABC1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 02-Feb-2001
C:Accession: A54774
R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A:Reference number: A54774; MUID:94375008; PMID:9088782
A:Accession: A54774
A:Molecule type: mRNA
A:Residues: 1-2201 <LUC>
A:Cross-references: GB:X75926; NID:9495256; PIDN:CA53530.1; PID:9495257
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop
F:856-1047/Domain: ATP-binding cassette homology <ABC1>
F:873-880/Region: nucleotide-binding motif A (P-loop)
F:1869-2060/Domain: ATP-binding cassette homology <ABC2>
F:1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match	49.6%;	Score	5531.5;	DB 2;	Length	2201;
Best Local Similarity	49.4%;	Pred. No.	0;			
Matches	1107;	Conservative	363;	Mismatches	574;	Indels 195; Gaps 27;
Qy	61	LPSAGTVWLOGLICNVNNTCFPOLTGCEPGRSLNFNDLSVSRLLADARTVLGGASAH	120			
Db	1	MPSAGTLPWQGIICNANNPCFRYPPTGEAPGVVGNFNKSIIVSRFLFSDAQRLLLSQRDT	60			
Qy	121	TLAGLGKLIATLRAAR-----STAQPQPTKQSPLEPPM-----	153			
Db	61	SIKMHKVLRLMRQIKHPNSNLKQDFLVNDNTEGFLQHLNLSLPRSTVDSLLQXNVGLQ	120			
Qy	154	-----LDVAELL-----	160			
Db	121	KVFLQGYQLHLASLNGSKLEBIIQLGDAEVSALCGLPRKKLDAARVRLRYNMDILKPVV	180			
Qy	161	-----TSLRTESLGLALGOAQEPHLSLEAAEDLAQELLALRSVLER-----ALLQRP	212			
Db	181	TKLNSITSLHPTQHLA-----EATTVLDSLGLGLAQLFSTKSWDSMRQRYMFTLVNS	233			
Qy	213	TSGPLEL-----LSALCSVRGSPSTVGPNSLNWYASDLMLVGV-----QRPESALPDSSLS	265			
Db	234	SSSSTQIYQAVSRIVCGHPGEGGLKIKSLNWDYNNYKALFGGNNTEDVDVTFYDNSTTP	293			
Qy	266	ACSELIGALDGHPLSRLLRWRLKPLIIGKLIAPDTPFTRKLMQAVNRTFEEITLLRDV	325			
Db	294	YCNLDLKNLESSPLRIIWKALKPLLVGKILVTPDTPATQVMAEVNKTQFLAVFHDLE	353			
Qy	326	EVEWMLGPRIITFMNDSSNVAMQLRLQMQ-----DEGRQRPGRGRDHMEALRSFL-----	377			
Db	354	GMWEELSPQIWTFMENSQEMDLVRTLDSRGNDQFWEQKLDGLDWTQAQIMAFKPNPD	413			

```
Qy 378 --DPGSGYQWDAHADVGHVGLGRVTECLSLDKLEAAPSEAAIVSRALQLLAEHRFW 435
Db 414 VQSPNGSVYTWREAFNETNQAIQIOTISRFMECNVNLKLEIPTTEVRLINKSMELDERKFW 473
Qy 436 AGVVFELGPEDSSDTEPTDPLG--PGHVRIRIKRMDIDVTRTNKIDRWDPCGPAADPL 493
Db 474 AGIVFTG-----ITPDSVELPHHVYKIKRMDIDNVERTNKIKDGYWDGPRADPF 523
Qy 494 TDLRYVMGGFYLLQDLVERAAVRLSGANPRAGLYLQOMPPYPCYVDDVFLRVLRSRSLPLF 553
Db 524 EDMRYVMGGFAYLQDVVEQAIIRVLTGSEKKTGYVVOQMPYPCYVDDIFLRVMRSRMLPF 583
Qy 554 LTLAWIYSVTLTVKAVVREKETRLDTRMRANGLGRAVLWGLFSLCGLPFLLSAALLVLV 613
Db 584 MTLAWIYSVAIIKSIIVYEKARLKTMRIMGLDNGILWFSWFSVSSILPLILVSAGLLWVI 643
Qy 614 LKLGDIPLYPHGPVGVFLPAAFAVATVYTSPLLSAPFSRANLAAACGLAYFSLYLPVL 673
Db 644 LKGNLLPYSDPSVVFVFLSVFAMVTILQCLFLSTLFSRANLAAACGGIYFLDYLFPVL 703
Qy 674 CVAWRDLPAGRVAAALLSPVAFPGCESLALLEQGEAGAHNVGTRPT-ADVFSLAQ 732
Db 704 CVAQDVVGFISIKIFASILLSPVAFGFCYFALPEEQIGVQWDLNLFESPVEEDGFNLTT 763
Qy 733 VSGLLLLDAALYGLATWYLEAVCGQGIPEPWNFPFRRSYWCQGRPPKSPAPCTPLD- 791
Db 764 AVSMMLFDTLYGVMTWIYIEAVFPQGIPIRPMPYFCTKSYWFGEEIDEKSHPGSSQGV 823
Qy 792 PKVLVEAPCLSPGVSVRSLEKEFPGPSQPALRGLSLDFVQGHITAFLGHNGAGKTTTL 851
Db 824 SEICMESEPTHLRGVSIQNLVKYRDGMKVADGLALNFYEQQITSLFLGHNGAGKTTTM 883
Qy 852 SILSGLFPFGGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTVDHWFVYGRLLKGL 911
Db 884 SILTGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHWFVYARLKGL 943
Qy 912 SAAVGPEDRLLODVLG--VSKQSVQTRHLSCGMQRKLSVAIAFVGGSQVWILDEPTAGV 970
Db 944 SEKHVKAEMEOMALDVLGPPSLKSKTSQLSGGGMQRKLSVALAFVGGSKVWILDEPTAGV 1003
Qy 971 DPASRRGIWELLKYRGRRTLILSTHLDRAELIGDRVAVVAGGRLCCGSPFLRRLHGL 1030
Db 1004 DPSRRGIWELLKYRGRRTIILSTHLDRAELIGDRVAVVAGGRLCCGSPFLRRLHGL 1063
Qy 1031 SGYYTLVLKARLPLTTN-----EKADTDMEGSVDTROEKNKGSGRSGVGTPL 1078
Db 1064 TGYTLVLKVDVSSLSRCSNSTSVSLCKEDSVSSQSDAGLGSDESHTLTDVSAI 1123
Qy 1079 LALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAELRLTGIGSDTS 1138
Db 1124 SNLIRKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRLDLGSISSYGETT 1183
Qy 1139 LEEIFLKVVECAADTDMEDG-----SCGOHLCTGAGLVDLRLKMPQ 1183
Db 1184 LEEIFLKVVEGSDAETSOGTLTPARRNRARFGDKQC-LHPFTDDAVID-----PN 1234
Qy 1184 ETALENGEPAGSAPETDQGGSDAVG--RVQGWALTQQQLQALILLKRLFLARSRRLGFA 1241
Db 1235 DSDID-----PESRETLLSGMDCKGSYQLKGWKLTOQQFVALLWKLLIARRSRKGFFA 1289
Qy 1242 QIVLPALFVGLALVFSILVPPFGHYPALRLSPTWYGAQVFFSFSDADGDPGRARLLBALL 1301
Db 1290 QIVLPAVFVCTALVFSILVPPFGKYPSLELQPMWYNEQYTFVNSNADPEDMGTOELLNALT 1349
Qy 1302 QEAG-----LEPPVQVHSHRFSAPVPAEVAKVLASGNWTPESPSPACQSQPG 1351
Db 1350 KDPGFGTRCMEGNDIPDTPCLAGEDWTISPVPOSIVDLFQNGWNTWKNFSPACQSSDK 1409
Qy 1352 ARRLPDCPAAAGGPPPPQAVGTSGEVQVNLGTGRNLSDFLVKTYPRVLVRGLKTKKWNE 1411
Db 1410 IKXMLVPCPGAGGLPPFPQRKQTADILQNLGTGRNISDLVKTYYQIIAKSLKMKIWNNE 1469
```

```
Qy 1412 VRYGGESLG--GRDGLPSGOELGRSVEELWALLSPLPGGALDRVLKVLNTAWAHSLDQADS 1470
Db 1470 FRYGGESLGVSNSQALPPPSHEVNDAIKQMKKLLKLTKTDSADRPLSSLSGRFMAGLDTKNN 1529
Qy 1471 LKTFWNNKGHWSVAFVWNRASNAIHLRAHLPPGPARHAHSITTLNHPNLNLTKQOLSEALM 1530
Db 1530 VKWFWNNKGHWAISSEFLNVNNAILRANLQKGENPSQYGITAFNHPNLNLTKQOLSEVALM 1589
Qy 1531 ASSVDVLVSICVVPFAMSFVPASFTLVLIBERVTRAKHLQMLGGLSPTLYWLNFLWMCN 1590
Db 1590 TTSVDVLVSIICVVPFAMSFVPASFTLVLIBERVTRAKHLQMLGGLSPTLYWLNFLWMCN 1649
Qy 1591 YLVPACTVWLIFLAFQORAVVAPANLPAIIILLLLCWMSITPLMYPASFPFSPSTAYVV 1650
Db 1650 YVVPATLVIIIFICFOOKSYVSTNLPVLAIIILLLLYGWSITPLMYPASFPFSPSTAYVV 1709
Qy 1651 LTCINLPIGINGSWATEVLELFSQDKQVSRILKQVFLFPFPCGLRGILDMVRNOAMA 1710
Db 1710 LTSVNLFIGINGSVAIFVLELFTNNKLNINDILKSVFLFPFPCGLRGILDMVRNOAMA 1769
Qy 1711 DAFERLGDROFQSPLRVWVGKNNLWVIOGPIFLFLFTLLQHRSQLLPQPRVRSRLPLG 1770
Db 1770 DALERFGENRFVSPSLWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVKAKLPPLN 1829
Qy 1771 EEDVYAREERVVQATQGDVLVRLNLTKVYRGQRMPPAVDRCLGIPGCEGCLLVNG 1830
Db 1830 DEDEDVRRERQRIIDGGQNDIILEIKELTKIYRKRKPPAVDRICIGIPPGCEGCLLVNG 1889
Qy 1831 AGKTSIPRMVGTDTLASRGEAVLAGHSVAREPSAAHLSMGVYCPSQSDAIFELLITGREHLEL 1890
Db 1890 AGKSTTFKMLTGTPTVTRGDAPFLNKSILSNIHEVHQMGYCPQFDAITELLITGREHVEF 1949
Qy 1891 LARLGVPEAOVAQTAGSLARLGLSWYADRPAGTVSGGNKRKLATALALVGDPAVVFLD 1950
Db 1950 FALLRGVPEKEVGKFGEWARKLGLVKYGEKYSASNYSGGNKRKLSTAMALIGGPPVVFLD 2009
Qy 1951 EPTTGMDPSARRFLWNSLLAVREGSRVMLTSHSMECEALCSRLLATMVNGRERCLGSPQ 2010
Db 2010 EPTTGMDPKARRFLWNCALSIKVEGRSVLTSMSMECEALCTRMALMVNGRFRCLGSPQ 2069
Qy 2011 HLKGRFAAGHTLRLVPAARS--QPAAAFVAASFPGSELREAHGGRRLRFPPOGGRCALA 2068
Db 2070 HLKXNRFQDGTIVVTRAGSNPDLKPQVEFFGLAFPGSVLKEKHNNMLOYQL-PSSLSLA 2128
Qy 2069 RVFGEIAVHGAHGVEDFVSQTNLEBFLVYFKDQKQDE---DTEQKEAGVGVDPAPG 2125
Db 2129 RIFSILLSQSKRLHIEDYSVSQTTLDQVFNFAKDQSDDDHLKDLSLHKQTV-VDVAV- 2186
Qy 2126 LQHPKRVSQFLDDPSTAET 2144
Db 2187 -----LTSFLQDEKVKES 2199
```

RESULT 2

A59189

Atp-binding cassette transporter - human (fragment)

N/Alternate names: KIAA1062 protein

C/Species: Homo sapiens (man)

C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004

C/Accession: A59189

R/Kikuno, R.; Nagase, T.; Ishikawa, K.; Hirose, M.; Miyajima, N.; Tanaka, A.; Kotani,

DNA Res. 6, 197-205, 1999

A/Title: Prediction of the coding sequences of unidentified human genes. XIV. The comple

A/Reference number: Z22961; MUID:9937452; PMID:10470851

A/Accession: A59189

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-1529 <KIK>

A/Cross-references: UNIPROT:Q9HC28; GB:AB028985; NID:g5689460; PIDN:BAA83014.1; PID:d104

A/Experimental source: chromosome 9; clone hJ03579; clone lib paluescriptii SK plus; tis

C/Genetics:

A/Map position: 9

A/Note: KIAA1062

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

```
Query Match      25.7%; Score 2863.5; DB 2; Length 1529;
Best Local Similarity 41.1%; Pred. No. 3.7e-171;
Matches 644; Conservative 246; Mismatches 438; Indels 239; Gaps 36;

QY 745 GLATWYLEAVCGQGIPEPWPFPFRSVCWGP-RPPKSPAPCPPLDPKVLV-----796
D 1 GILTWYIEAVHPCMTGLPRWTFPLQKSYWLGSGRTEAWESWPWARTPLRSVMEEDQAC 60

QY 797 -----EAPPGCLSGVSVRSLEKFPSPQALRGISLDFYQGHITAFLGN 843
D 61 AMESRRFESTRGMEBEPTLPLVVCVDKLTQVYKDDKLAANKLSNLNYENQVVSFLGN 120

QY 844 GAGKTTTSLSGLPSPGSGSAFILGHVDRSSMAAIRPHLGVCPQVNVLFDMLTVDHVV 903
D 121 GAGKTTTMSILGFPPTSGSATIYGHDIRTEMDEIRKNLGNCQHNVLFDRLTVEEHL 180

QY 904 FYGRKLGLSAAVVGPEQDRLLQDVLGVSKQSVQTHLSGGMQRKLSVAIAFYGGQVTL 963
D 181 FYSRLKMAQERIRREMDKWIEDLSNKRHSLVQTLSGGMRKLSVAIAFYGGSAIIL 240

QY 964 DEPTAGVDPASRGITWELLKYREGRTLLSTHLLDEALLGDRVAVVAGGRLCCGSP 1023
D 241 DEPTAGVDPYARRAIWDLTKYKRTILLSTHLLDEALLGDRVAVVAGGRLCCGSP 300

QY 1024 FLRRHLGSGYLLTVKARLPLTNEKADTDMGSGVDTRQEKNGSGSGRVGTPQLLALVQ 1083
D 301 FLKGYGGRYLLTVKR--PAEPGGFQEPGLASSPPGRAPLSSCS-----LQVSQIR 352

QY 1084 HWVPGARLVEELPHELVLVLPYTGADHGSFATLRELDRALBELRLTGVGSDTSLERIF 1143
D 353 KHVASCLLVSDTSTELSYLPEAAKKAFAERLFQHLERSLDALHLSPLGLMDTLEEVF 412

QY 1144 LKVVPE-----CAADT-----DMEDGSGQHLCTGIAGLDVTLRKMPQPTALENGEP 1192
D 413 LKVSSEDSLENSADVRESKRDVLPFAGEPASGEGHAG-NLARCSELTQSQASLSQAS 471

QY 1193 AGSAPETDQSG-----PDVAVG--RVQGWALTR-----1218
D 472 VGSA--RGDEGAGYTDVYGDYRPLFDNPQDPDNVSLQVEAEALS RVQGSRLKDGWLKV 530

QY 1219 QQLQALLKRFLLARRSRGLPAQIVLPALFVGLALVSLIVPPFGHPYALRLSTMY--1276
D 531 RQFGLLVKRFHFCARRSKALFSQLLPALFAFFCVAMTVALSVEITGDLPLVLSPSQVHN 590

QY 1277 -----GAQVSFSED-----APGDPGRALLLEALLQEAG-----LEEP-----PV 1311
D 591 YTPQGRNFIPYANERREYRLRLSP-DASPOQLVSTFRLPSGVGATCVLKSFGANGSLGPT 649

QY 1312 QHSSH-----RFSAPVPAEVAKVLAGSNWTPESPSPA-----1344
D 650 LNLSSGESRLLAARFFDSKCLSFSTQGLPLSNFVPPPPSPAPSDSPASDEDLQAWNVL 709

QY 1345 -----COCQOPGARRLLPDCPAAAGPPPPPPQAVTGSGEVVQ 1380
D 710 PPTAGPEMWTSAPLRLVREPVRCCTCAQGTGF---SCPSVSGHGPQMRVV-TGDLT 765

QY 1381 NLTGNSLSDFLVKTYPRLVRQGLTKKWNVREYVGFSLGGNDPGLPSGQLGRSVBELW 1440
D 766 DITGHNVSEYLLFTSDRFLH-----RYGAIITFG-----794

QY 1441 ALLSPLPGGALDRVLKNTA-----WAHSLDAQDSLKTIWNNKWSMVAFVNRASN 1492
D 795 -----NVLKSIPASFGTRAPPMWRKIAVRRAAQVFNKNGHSMPTYLSLNN 842

QY 1493 AILRAHLPP-----GPARHAHSITTLNHLPLNLTKEQLESAALMASSVDVLVSCVWFAMSFV 1549
D 843 AILRANLPKSKGNPA--AYGITVTHNPWNKTSASLSLOYLL-QGTDVVIATFIIVAMSFV 899

QY 1550 PASFTVLVIERVTAKHQLMGLGSLPTLYWLGFLNDMCMVLYPACIVLWIFLAFQORA 1609
D 900 PASFVYVFLVAEKSTAKRKHQFVSGCNPIIYWLANYVMDMLNLYLVPATCCVILFVFDLPA 959
```

```
QY 1610 YVAPANLPALLLLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMATFVL 1669
D 960 YTSPTNFPAVLSLFLYLGWSITPIMYPASFWEVPSAYVFLVINLFIGITATVATFLL 1019

QY 1670 ELFS--DQKLEQVSRILKQVFLIPHPCLGRGLIDMVNRNQAMADAFERLGD-RQFQSPLRW 1727
D 1020 QLFEHDKDLKVNSYLKSCFLIFPNYNLGHGLMAYNEYINNEYAKVQFQDKMKSPFEW 1079

QY 1728 EVVGKNLAMWTQGPLFLLFTLLLOH-----RSQLLPQPRVRSLSPLLGEDEDEVARRERV 1783
D 1080 DIVTRGLVAMAVEGVVGGELLTIMCQYNFLRRPQRM---VSTKPV--EDVDVASSRQVR 1134

QY 1784 VQAGTQGVDLVLRNLTKVYRQO---RMPADVRLCLGIPGEGCEGGLGVNGAGKTSFRMV 1840
D 1135 LRGDADNDMWKLENLTKYKSRKIGRILAVDRCLGVRRPGECEGGLGVNGAGKTSFVKL 1194

QY 1841 TGDITLASGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELITGRELHLLARLRGVRPA 1900
D 1195 TGDESTTGGEAFVNGHSLKELLQVQSLGYCPQCDALFDELTAAREHLQLYTLRLGISWK 1254

QY 1901 QVAQTAGSLARLGLSWADRPAGTYSNGNKKLATALAVGDPVAVVFLDEPTTGMDPSA 1960
D 1255 DEARVVVKWALEKLELTUYADKPAGTYSNGNKKLSTALIGYPAFIFLDEPTTGMDPKA 1314

QY 1961 RRFLMNSLLAVVREGRSVMLTSHSMECEALCSRLAIVMNGFRCLGSPQHLKGRFAAGH 2020
D 1315 RRFLMNLILDLIKGTSVVLTSMSMECEALCRLAIVMNGRLRCLGSLQHLKNRFGDY 1374

QY 2021 TLTLRVPAARS--QPAFAVFAAEFPGSELREAHGRLRLQLPGGRCALARVGEELAVHGA 2079
D 1375 MITVTKSSQSVKDVVRFPNRFNPEAMLKRHTTKVQYQL-KSEHISLAQVFSKMEQVSG 1433

QY 2080 EHGVEDFSVQTMLEEVFLYPSKDGKDBETEOKEAGVGVDPAQGLQHPKRVQSFLDDP 2139
D 1434 VLGIEDYSVQTMLEEVFLYPSKDGKDBETEOKEAGVGVDPAQGLQHPKRVQSFLDDP 2139

QY 2140 STAETVL 2146
D 1487 RSAPTEL 1493

RESULT 3
BS4774
A: Species: Mus musculus (house mouse)
C: Species: Mus musculus (house mouse)
C: Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 16-Aug-2004
R: Accession: BS4774
R: Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A: Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A: Reference number: A54774; MUID: 94375008; PMID: 8088782
A: Accession: BS4774
A: Molecule type: mRNA
A: Residues: 1-1472 <LUC>
A: Cross-references: GB:X75927; NID:g495258; PIDN:CAA53531.1; PID:g495259
C: Superfamily: ATP-binding cassette homology
C: Keywords: ATP; nucleotide binding; P-loop
F: 44-234/Domain: ATP-binding cassette homology <ABC1>
F: 161-68/Region: nucleotide-binding motif A (P-loop)
F: 1108-1300/Domain: ATP-binding cassette homology <ABC2>
F: 1126-1133/Region: nucleotide-binding motif A (P-loop)

Query Match      24.2%; Score 2694.5; DB 2; Length 1472;
Best Local Similarity 40.8%; Pred. No. 1.4e-160;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

QY 796 VEEAPPGLSPGVSVRSLEKFRFPSPQALRGISLDFYQGHITAFLGHNGAGKTTLSLS 855
D 16 MEEEPHPLPLVVCVDKLTQVYKNDKLAANKLSNLNYENQVVSFLGHNGAGKTTNSILT 75

QY 856 GLFPPSGSAPILGHDRVSSMAAIRPHLGVCPQVNVLFDMLTVDHVVYFGLKGLSAAV 915
D 16 MEEEPHPLPLVVCVDKLTQVYKNDKLAANKLSNLNYENQVVSFLGHNGAGKTTNSILT 75
```


QY 307 LMAQVNRTEBELTLRLDRVREVMELGPRIFFTEFMDSSNVAMLRQLLQMDQERRPRPG 366
D 9 LLLKXNYTLQKRVLVLELPL--PLLF-----SGILLRLKIQSENVFNATYPG- 59
QY 367 RDHMEALSRFLDPGSGGYSWQ---DAHADVHLVGTGLGRVTECLSLDKLEAPSEALV 422
D 60 -QSIQELPLFFTPPGDTEWELAYTPSHSDAAKTVTETVRRAALVNM-RVRGFPSEKDFE 117
QY 423 SRALQLLAEBHRWAGVVFGLGPDSDPTBHPPTDLGPGHVRIRKIMDIDVTR----- 475
D 118 DYIRDNCSSSVLAADVFEHPNHS---KEPLPLAVKHLRFSYTRNMTQTSFFLK 174
QY 476 -----TNKIRDFWDGPAADPLT-----DLRYVMGGFYVLQDLVERAAVRVLSGAN--- 522
D 175 ETEGWHITSLFLFPNPGP-RETPSPDGEGPCYIREGLAVQHAVDRAIMEYHADAATRO 233
QY 523 --PRAGLYLOQMPYCYVDVDFLRLVLSRLPLFLTLAMTYSVTLTKAVRKEKTRLD 580
D 234 LQRLTVTIKRPFPPTADPLFAIQQLPLLLLLSFTYTALTITARAVVQEKRLKEY 293
QY 581 MRAMGLSRAVLWGLNFLSCGCLGPFLLSALLVLVLKLG-----DILPYSHPGVVFLFLAAF 635
D 294 MRMGLSSWLHWSANFLFLFLLLTAASPTWLLFCVKVKNVAVLSRSDPSLVLAFLCF 353
QY 636 AVATVTSQFLLSAFFSRANLAAACGLAYFSLYPLVLCVAVWRDLRDPAGRVAASLLSPV 695
D 354 AISTISFSPFVSTFSKANMAAAGFLYFTYIYFVFVAPRYNMWTLQKLSCLLSNV 413
QY 696 AFGFCESLALLEBOGEGQWNVTRPTA-DVFSLAQVSGLLLDAAALGLATWLEAV 754
D 414 ANAMGAQLIKFPAKGMGIQWRDLSPVNVDDDFCGQVLGMLLDLSVLYGLVTVMEAV 473
QY 755 CGQVIGIPEPMNFPFRSRVWCQPRPKSPACPPTPLDPKVLV-----REAPGLSPGVSV 809
D 474 FPGQGVQPWTFITMPSWCS-KERAVAGKEEDSDPEKALURNEYFAEPEDLVAGIKI 532
QY 810 RSLEK--RFPQSPQALRGLSLDFYQGHITAFILGHNGAGKTTLSILSGLFPPSGSAPF 867
D 533 KHLKSVFRVGNKDRAVRDLNLNLEGGQITVLGHNGAGKTTLSMLTGLFPPTSGRAVI 592
QY 868 LGHDVRSMAAIRPHLGVCPQYVNLFDMLTVDEHWFVGRGLKLSAAVVGPDRLQDV 927
D 593 SGYEISQDMVQIRKSLGLCPQHDILFDNLVVAEHLFYAQLKGLSRQCPPEVKOMLHI 652
QY 928 GLVSKQSVQTRHLSGMOKLSVATAFVGGSOVWTLDEPTAGVDPAARRGIWELLKYRE 987
D 653 GLEDKNSRSRFLSGMRKLSIGALTAGSKVLILDEPTSGMDAISRAIWDLLQROKS 712
QY 988 QRTILSTHLLDEALLGDRVAVVAGRLCCGSPFLFRRHLGSGYTLTKARLPLATTN 1047
D 713 DRTIVLTTHFMDADLLGDRIAMAKGELQCCGSSFLKQKYGAGYHMTLVKE----- 765
QY 1048 EKADTDMGSDVTREKNGSGSRVGTFTQALLVQHVWPGARLVEELPHELVLVLPYTG 1107
D 766 -----PHCNPEDISQLVHHVHPNATLESSAGAELSFILPRES 802
QY 1108 AHGDSFATLFRDLTRLAELRTGYSIDTSEELFLKV----- 1146
D 803 TH--RFEGLFAKLEKQKELGASFGASITWEEVFLRVGKLVDSMDIQAIQLPALQYQ 860
QY 1147 VEECAADTDMEDSGCGHLLCTGTGIAGLDVTLRLKMPPEQTALENGRSPAGSAPETDQSGPD 1206
D 861 HERRASDWAVDNLCG-----AMDPSDGIALLIEER---T 893
QY 1207 AVGRVOGNALTRQQLQALLKRLFLARSRRGLFQAIVLPAFLVGLAVLSLVPFGHY 1266
D 894 AVKLTGLALHCCQFWAMFLKKAASWREKMAVAQVLVPLTCVTLALLAINYSSELFPDD 953
QY 1267 PALRLSPMTYGAQVQSFEDAPGDCRARILLALLQEAAGLEPPVQHSRHPFAEVPFAE 1326
D 954 PMLRLTLGEYGRVTVFVSF-----FQTSQGOQL-----SEH----- 985

QY 1327 VAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGEVVQNLGTRN 1386
D 986 -----LKDALQAG--QSPREVLG-----D 1003
QY 1387 LSDFLVKTYPRLVROGLTKKVVNEVRYGGFSGLGRDPLGSGOELGRSVBELWALLSPL 1446
D 1004 LBEFLI-----FRASVEGGFN----- 1020
QY 1447 PGAGLDRLVKNLTAWAHSIDAQDSLKIWFNNKGMHSMVAFVNRASNAIRLAHLPPGPARH 1506
D 1021 -----ERCL--VAASFRDUGERTVYNALFNQYHSPATALAVDNLKFK--LLCGP--H 1069
QY 1507 AHSITTLNHP-----LNLTKQBSAALMASSVDVLVSVVFAFMSFVPASTFLVIER 1561
D 1070 A-SIVVSNFPQPSRALQAQKQFNEG---RKGFIDALNL--LFAMAFLASTESILAVSER 1123
QY 1562 VTRAGHLQMGSLFTLWLGFLWDMCNLYLPACIVLILFLAFOORAVVAPANLPALLL 1621
D 1124 AVQAKHVQFVSGVHVASFMLSALLWDLISFLPSLLLVVFKAFDVRAFTRDGHMADTLL 1183
QY 1622 LLLLYGWSITPLMYPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLEL---FSDQKIQ 1678
D 1184 LLLLYGNAIIPLAYLWNEFFFLGAATAYTRLTIFNLISGI---ATFLMTIMRIAPVAKLE 1239
QY 1679 EVSRILKQVFLIPFHCILGRGLIDMVN-----QAMADAFERLGRDQFQSP-LRW 1728
D 1240 ELSKTLDHVFLVPLNHLGMAVSSPYENVETRRYTSSEVAHCKYKNIQIENFYAWS 1299
QY 1729 V--VGNLLAMVIQ-----PLFLFTLLQHSOLLPOQVR-----SLPLLGEED 1773
D 1300 APGVGRFVMAAASCAYLILFLITETNLLQRLGILCALRRRTLTETYTRMPVL-PED 1358
QY 1774 EDVARERERVQGATQG---DVLVLNLTKVYRGORMP--AVDRLCIGIPPGCEGGLGV 1828
D 1359 QDVADERTRIIAPSPDSLLHTPLIIRKLSKVTE-ORVPLLAVDRLSLAVQKCEGGLGF 1417
QY 1829 NGAKTSTFRMVTGDTLASGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHL 1888
D 1418 NGAKTSTFRMVTGDTLASGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHL 1888
QY 1889 ELLARLGRVPEAQVATAGSLARLGLSWYADRPACTYSGGNKRKLATLALVGDPAVVP 1948
D 1478 VMYARLGRIPERHIGACVENTLRGLLLEPHANKLVRTYSGGNKRKLSTGIALIGSPAVIF 1537
QY 1949 LDEPTTGMDSARPLWNSLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGS 2008
D 1538 LDEPTTGMDSARPLWNSLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGS 2008
QY 2009 FOHLKGRFAAGHTLTLPAAARSQPA-----AAFVAAEFFGSRELAHGRRLRFQLPPGGR 2064
D 1598 FOHLKSGFGSVSLRAKVSQSQQALBEFAFVDLTFPGSVLEDEHQGMVHYHL-PGRD 1656
QY 2065 CALARVFGELAVHGAHGVDFSVSOTMLEEVFLYFS 2101
D 1657 LSWAKVFGILKAKKKEKYGVDDYSVQSILEQVFLSFA 1693

RESULT 5

A59188

ATP-binding cassette transporter ABC3 - human

C:Species: Homo sapiens (man)

C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004

C:Accession: A59188

R:Conors, T.D.; van Raay, T.J.; Petry, L.R.; Klinger, K.W.; Landes, G.M.; Burn, T.C.

Genomics 39, 231-234, 1997

A>Title: The cloning of a human ABC gene (ABC3) mapping to chromosome 16p13.3.

A:Reference number: A59188; MUID:97179225; PMID:9027511

A:Accession: A59188

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1704 <CON>

A:Cross-references: UNIPROT:Q99758; GB:U78735; NID:g1699037; PIDN:AAC50967.1; PID:g16990

C:Genetics:

A;Gene: GDB:ABC3
A;Cross-references: GDB:3770735; OMIM:601615
A;Map position: 16p13.3-16p13.3
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

	Query Match	22.5%;	Score 2502;	DB 2;	Length 1704;
	Best Local Similarity	33.7%;	Pred. No. 2.2e-148;		
	Matches 639;	Conservative 309;	Mismatches 635;	Indels 314;	Gaps 47;
	C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology				
Qy	307	LMAQVNRITFEELTLRLDRREVW-EMLGPRIFTFMNDSSNVAMLQRLQLMQDEGRQPRPG	365		
Db	9	LLLWKNYTLQREKVLTVLEFLPLFPGLIW-----URLKIQSENVNATYPG	59		
Qy	366	GRDHMEALRSFLDPGSGGYSWQ-----DAHADVGLHVTGLGRVTECSLSDKLEAPSPAAL	421		
Db	60	--QSIQELPLFTFTPPPGDTWELAYIPSHSDAAKTVTETVRRALVINM-RVRGFPSEKDF	116		
Qy	422	VSRALQLLAEHRFWAGVVFGLGEDSDPTHTPDLGPGHVRKIRKMDIDVTRTKNI--	479		
Db	117	EDYTRYDNCSSVLAAVVEHFHNHS---KEPLPLAVKYHLRFSYTRNNYMTQTSGFFL	173		
Qy	480	-RDRFWD-----PGPAADPLT-----DLRYVMGFGVYLQDLVERAAVRVLSGAN---	522		
Db	174	KETECWHITSLPLFPNCPRELTPDGPCEPVIREGFLAVQHADVRAIMEYHADATRQ	233		
Qy	523	--PRAGLYLQOMPXYCYDDVFLRVLRSRLPLFTLAWIYSVTLTVKAVVREKETRLURDT	580		
Db	234	LFQRLTVTIKRPYPYPFIADPELVAIQQLPLLLLSFTYLTALTIAAAVVOEKERRLKEY	293		
Qy	581	MRMGLSRAVLWLGLFSLCLGPFLLSAAALLVLVLKG-----DILPYSHPGVVFLPLAAF	635		
Db	294	MRMGLSLGWLHWSAWFLFFFLFLLIAASFTMLLFCVKVPKNVAVLSRSDPSLVLAFLLCF	353		
Qy	636	AVATVQTQSLFSAFRRANLAAACGLAYFSLYLPVVLVAVMRDLRPLGACGRVAASLILSPV	695		
Db	354	AISTISFSFMVSTFTFSKANMAAAFGGLFYFTYIPYFVAPRYNNMTLSQKLCCLLSNV	413		
Qy	696	AFGFGCESLALLEBOGEGAWINVGTRPTA-DVFSLAQVSGLLLLLDAALYGLATWYLEAV	754		
Db	414	AMAMGAQIIGKFEAKGMGIQWMDLLSPVNVDDFCFGQVLGMLLLDSVLYGLVWTYMEAV	473		
Qy	755	CPGOVGIPPEPWNFPFRSSYWCQPRPKSPAPCTPLDPPKLV-----EEAPGLSPGVSV	809		
Db	474	FPQGFGVQPPYVFFIMPSWVG- KPRAVAKGKEEDSDPEKALRNEFEAPEDLVAGIKI	532		
Qy	810	RSLEK--RFGQSPQALRGLSLDFYOGHITAFGLHNGAGKTTLSILSGLFPFSGGSAFI	867		
Db	533	KHLSKVFVRGNKDRAAVRDLNLNLYEGQITVLLGHNGAGKTTTSLMLTGLFPPTSGRAYI	592		
Qy	868	LGHVRSMAAIRPHLGVCPQNVFLFDMLTVDHFWFYGLKGLSAAVGVPEQDRLLQDV	927		
Db	593	SGYEISQDMVQIRKSLGLCPQHDILFDNLTVAEHLFYAQLKGLSKQKCEPVKQMLHII	652		
Qy	928	GLVSKQSQVTRHLSCGMORLKVAVAFVGSQVILDEPTAGVDPASRRGIGWELLKLYRE	987		
Db	653	GLEDKNGSRRLSGCMRKLIGIALIAGSKVLIIDEPSTGMDAISRRAINDLLQOKS	712		
Qy	988	GRTLTLSTHLLDEALLGDRVAVWAGGRLLCCGSPFLFRLRHLGSGYYTLTKVARLPLTTN	1047		
Db	713	DRITVLTTHFMDADLLGDRIAIMAKGELQCCGSSLFLKQKYGAGYHMTLVKE-----	765		
Qy	1048	EKAUTDMEGSVTRQEKNGSQSRVGTQQLLALVOHWVPGARLVEELPHEILVLVLPYTG	1107		
Db	766	-----PHCPNBDISQLVHHVVPNATLESSAGAELSFILPRES	802		
Qy	1108	AHDGFSATLFRDLTRLABELRTGVGISDTSLEEEIFKV-----	1146		
Db	803	TH--RFEGLFAKLEKKQKELGTASFGASITTWEEVFLVRGKLVDSMDIQAIQLPALQYQ	860		
Qy	1147	VEECAADTMDGSCQHQHICTGIAGLDVTLRLKMPPOETALENGEPAGSAPETDQSGSPD	1206		
Db	861	HERRASDWAVDSNLCG-----AMDPSDGTGALIEBER---T	893		

Qy	1207	AVCRVQGWALTROOLQALLIKRELLARRSRGILFAQIVLPALEFVGLALVPSLIVPPFGHY	1267
Db	894	AVKLNTGLALHCOQFAMFELKKAAYGWRERKMVAQVLVELTCTVTUALCALTAINSYSELFDD	953
Qy	1267	PALRLSPMYGAOVSYFSESDAPGDCGRARLEALLQEAUGLEEPVQHSHSRFSAPEVPAE	1326
Db	954	PMURLTIGEYGRVVPFSV-----PGTSJLQOQL-----SEH-----	985
Qy	1327	VAKVLASGNWTPSPSPACQOCQSGARRLLPDCPAAAGGPPPPQAVTGSGEVVQNLTGRN	1386
Db	986	-----LKDALQAEG--QEPREVLG-----D	1003
Qy	1387	LSDFLVKTYPRLVROGLKTKKWNVEYVYGFSGIGDRDGLPLSGQELGRSVEELWALLSPL	1446
Db	1004	LEBFLLI-----FRASVEGGGFN-----	1020
Qy	1447	PGGALDRVLKNTAWAHSLDAQDSLKTWFNNKGMHSMVAFVNRSNAILRAHLPPGPARH	1506
Db	1021	-----ERCL--VAASPRDVGERTVVNALFNQAYHSPATALAVVDNLFFK--LLCGP--H	1069
Qy	1507	AHSITTILNHP-----LNLTKEOJSEALMASSVDVLVSIICVWFAMSFVPASFTVLITIEER	1561
Db	1070	A-SIVVSNFPQPSALQAAKQFNEG--RKGFDIALNL--LFAMAFLASTFSILAVSER	1123
Qy	1562	VTRAKHLQLMGGLSPTLYMIGNFLMDCMNVLPACIVLVILFAPQORAYVAPANLPALLL	1621
Db	1124	AVQAKHVQFVSGVHVHASFMLSALLIWLISLILPSLILLVVFKAFDVRAFTRDGHMADTLL	1183
Qy	1622	LLLLYGWSITPLMYPASFFSFVSTAVVLTCTINLFIGINGSWATFVLEL--FSDOKLQ	1678
Db	1184	LLLLYGWAIPLMYLWNFPFLGGAATATYRUIFNILSGI-----ATFLWVTIMRIPAVKLE	1239
Qy	1679	EVSRILKQVFLIPPHFCFLGRLMDVRN-----QAMADAFERLGDGRQFOSPL-RWE	1728
Db	1240	ELSKTLDHFLVLPHNCLGMVASFVYENYETRYCTTSSEVAHYCKYKNIQYQENFYAWS	1299
Qy	1729	V--VGNKLLAMWTQ-----PLFLFTLLLOHRSOLLPPQPRV-----SLPLGEED	1773
Db	1300	APGVGRFVSAASGAYLILFLELITNLLQRLGLCALRRRTLTETLYTRMPVL-PED	1358
Qy	1774	EDVARRERVVQATQG--DVLVRLNLTKVYRGQMP--AVRCLCIGIPPGCFGLGV	1828
Db	1359	QDVADERTRILAPSPSLHTPTLIILKELSKYB-QRVPLLAVDRLSLAVOKGCEFGULLGF	1417
Qy	1829	NGAGKTTFRMVTGDTILASGEAVLACHSVAREFSAHLSMGYCPOSDAIFELLTGREHL	1888
Db	1418	NGAGKTTTFKMLTGBESLTSGDAFVGGHRISDVGVKVRQIRIGYCPQCPDALLDHMTGREML	1477
Qy	1889	ELLARLGVPEAOVQATAGGLARLGLSWADRPAGTYSGCNKRKLATALALVGDPAVVF	1948
Db	1478	VMYARLRGIPERHIGACVENTLRGLLEPHANKLVRITYSGNKRKLSTGIALIGEPAVIF	1537
Qy	1949	LDEPTTGMDSARRFLWNLSLLAVVRGRSVMLTSHSMEECEALCSRLAIWNGFRFCIGS	2008
Db	1538	LDEPSTGMDPVARLLMDVTVARARESGKALITISHSMEECEALCTRLAIWVQGFKCLGS	1597
Qy	2009	POHLKGRFAAGHTLTLRVPAAQSQA-----AAFVAAEFFPGSELREAHGGRRLRFOLPPGGR	2064
Db	1598	PQHLKSKFGSGYSIRAKVQSEGOQEALEEFKAFVDLTFFPSVLEDEHQGVHYYHL-FGRD	1656
Qy	2065	CALARVPEGLAVHGAHBGVDFSVQSTMLBEVFLYFS	2101
Db	1657	LSNAKVFGLLEKAKEKYGVDDYSGVQISLSQVFLSFA	1693

RESULT 6
T33783
hypothetical protein Y39D8C.1 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 16-Aug-2004
C/Accession: T33783
R/Becker, M.; Graves, T.; Yoakum, M.
submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid Y39D8C.

A:Reference number: Z21408

A:Accession: T33783

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1802 <BEC>

A:Cross-references: UNIPROT:Q9TXV8; EMBL:AF101313; PIDN:AAC69223.1; GSPDB:GN00023; CESP:

A:Experimental source: strain Bristol N2; clone Y39D8C

C:Genetics:

A:Gene: CESP:Y39D8C.1

A:Map position: 5

A:Introns: 45/3; 114/1; 195/1; 230/3; 543/3; 794/1; 849/1; 1036/2; 1099/1; 1132/3; 1165/

C:Superfamily: Atp-binding cassette homology

Query Match		16.4%; Score 1829; DB 2; Length 1802;
Best Local Similarity		28.4%; Pred. No. 4.2e-106;
Matches 539; Conservative 312; Mismatches 667; Indels 382; Gaps 55;		
QY	371	EMARFLDPGGGGYQWQAHADVHLVGLVTECLSLD---KLEAAPSEALYS---R 424
DB	124	DATQIMDKIQRYTANDLLNPAAYIKGLYNITAVPVLNTNNTYKGFTEGEMYSWMQG 183
QY	425	ALQLLAHRFAGVGVFLGPSSDTEHTPDGLGCHVRIKIRMDIDVVRN---KIR 480
DB	184	QFQSECDNPLLAGIVF---DSIAKDLKNPD-----KRDFTYIRLSNTHRSR 229
QY	481	DRFMDPGPAADPLTD--LRYVMG-----GFVYLQDLVERAAVRVLSGA 521
DB	230	NAFGDSVPWTSVSFAVQYVSGPINPDNDGSGPYQWEGMTVQRAVDVAITIIIGE 289
QY	522	N-----PRAGLY--LOOMPYPYVDDVFLRVLSRLSLFLTLAMYSVTLTVKAVVREKETR 576
DB	290	DAQLTPLDLSYQVSRFPFPGYSTKI-IBIGAFPMPIVIFSPMSTSVIIVRAVVVEKEDR 348
QY	577	LEDTRWRALSRVNLWGLFSLCLGPFLLSAALLVILKIGDILPYSPGVVFLAAPA 636
DB	349	LKEYNRVNLGSLQFINWVAHFI--INYAKLTFAVILVTILMHFVALKSDMTLMFVLMTYA 406
QY	637	VATVTSFLLSAFFSRANLAAACGLAYFSLYLPYVL--CVAMRDLPAAGRVAASLLSP 694
DB	407	FDVYVFAWISFPMNSATSATLISVFWMLLYFWYAFSSIDQTPYPLGRLINCINPD 466
QY	695	VAFGFCESLALLEQSGGAQHNVGTRPTAD---VFSLAQVSGLLLLDAALYGLATWYL 751
DB	467	IALNYGLQLAAAYETAQADGLKWGLFTPPSPDNNLTFGHALIA--LIVDGIIMILTWYI 524
QY	752	EAVCGQYGIPE-PNPFPRRSYWCGRPPPKSPAPCPTPLD-----PKVLEEAPPG 802
DB	525	EAVIPGGGVFPQKPFVYL-PSYWF-----PNSGSKTVDSDFQFOQIQYADHVKLEKEPTD 579
QY	803	LSPGVSVRSLEKRPFGS-----POPALRGLSLDFYOGHITAFIGHNAGKT 848
DB	580	LIPTINVNLTKYTSFPPKLFCKFKGSEKRAVSNLNLMYPPGQCTVLLGNAGKS 639
QY	849	TTLTSLGLFPFPGSGSAFILGHDVRSSMAAIRPHLGVCPQYNVLFDMLTVDHVMFYGLR 908
DB	640	TTFSMLTGVSFSGSAVYVNDIRTSPLKIRREMGLCPQYNTLFGFTVMEHLEFPAL 699
QY	909	KGLSAAVGPQDR--LLQDGLVSKSQSVQTHLSGGMOKLSVAIAFVGGSQVILLDEPT 967
DB	700	K---ERTWDPEAREILARLIDFKADPMAGALSGGQKRLSLAIALTIGSGSEVVMLEDEPT 756
QY	968	AGVDPASRRGWELLKYREGRTILSTHLDLAEILGDRVAVVAGGRLCCGSPFLRR 1027
DB	757	SGMPGARHETWLIQREKERTILLTFHMEADLLGDRIATWAGQLBCCGSPMLKQ 816
QY	1028	HLGSGYVTLVKARLPLTNEKADTMESGVYDTRQEKNGSGRGVGPFPQALLAQHVWP 1087
DB	817	QYGDGYHLTIV-----YDTTSPDVSKTTD-----IIREVIP 848
QY	1088	CARLVEELPHELVLVLPVTGAHDSFATLRELDTRIAEALRLTGVIGSDTSLBEIFLKVV 1147
DB	849	EAHVFSYIGQATYLLSAT--HRPIFPKLKLELHDHQTCCGITSFGVSIITMESEVFLKVG 906

QY	1148	BECAADTMEGSCGCHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGPD- 1206
DB	907	HTADERYVNEHG-----IEN--DISEMIEKDDPILQDL 937
QY	1207	-AVGRVQGWALTRQOLQALLKRFLARRSRRLGFAQIVLPALFVGLALVFSLIVFPFGH 1265
DB	938	RAQVRVTGFTLQMQHAKAMFYKRAIFFFRKWTQFLPQLVFPVAVLVLMVFTSQVLP---- 993
QY	1266	YPALRLSPMYCAQVSPFSEDAPGPGRGARLLEALLQEALEBPVQVSHSRFSAPDEVA 1325
DB	994	-----LGFVP-----SMFNFTSVSPSLKTFNFNFGLYTPALAITFTDSMILSQ- 997
QY	1326	EVAKVLASGNWTPESPSPACQCSQGARRLLPDCPAAAGGPPPPQAATGSGEVQNLTC- 1384
DB	998	-----ODPQTISLAPSDTKAG-----HLVSDSGNVTVLLGGS 1031
QY	1385	RNLSDFLVKTYPRL--VROGL-----KTKKWNVEYRGFSLGGRDPGLPSGOELGRSVEE 1438
DB	1032	QNLSSMVQGTVTQLGVTQTVVDITSNVEKFMIDQTN---AMGSRTFGLHYA----- 1079
QY	1439	LWALLSPPLPGGALDRVLKNLTAWAHSLOADSLKIWFNNKGHWSVAFVNRASNAILRAH 1498
DB	1080	-----LGFVP-----SMFNFTSVSPSLKTFNFNFGLYTPALAITFTDSMILSQ- 1123
QY	1499	LPPGPARHAHSITTLNHPL-----NLTKEQLSAAA--LMASSVDVLVSVVFAFMSFVP 1550
DB	1124	-----KQKQYQFTAVNHPLPSTQDTLKNTRSDGAAFLIAYGLLVISFVAVCV- 1171
QY	1551	ASFTLVLEERVTRAKHLQMLGGLSPTLYLNGFLNMDNCNVLVPAACIVVLIPLAFQORAY 1610
DB	1172	AGYSQFLITERKKSKHMQLLSGIRPWFMTAFIWDAAWFFIRILCFDAIFIEFNITAY 1231
QY	1611	VAPANLPAALLL--LLLYCWSITPLMPASFPSPSTAYVVLTCNLFIGINGSMAFVFL 1669
DB	1232	THDFGWMILITLFLLYGWTALPFTYFPQFFESAPKGFMMVMTVHILTMIGSIAVPTII 1291
QY	1670	E-----LFSDQKLOEVSRI-----LKVFLFPHP-----CUG 1697
DB	1292	SQTSLSLDAGYLSWIIFAWLFTYNIQIATVTFQNVNRIACKKLDCTIPMEKATACG 1351
QY	1698	RGLDMVRNQAMADAFERL-----GDROFQSPLRWEVVGKNNLAMVI-----QGFLPL 1745
DB	1352	-----TASERLYDNDVLFVGNR-----KGLVYVIFLAVQGFYV 1386
QY	1746	LFTLLQLHR--SOLLPPQVRSL--PLLGEDEEDVAREE-----RVQCATQGD 1791
DB	1387	IWVFMRENDQFTKLIALIRCKADNPIDITDTDKVDERVEDSDVIAEKSVMQRLANN 1446
QY	1792	--VLVRLNLTKVYRGQMPADVRLCLGTPPGCEGGLGVNGAGKTSTFRMTGDTLASRG 1849
DB	1447	KTALVSNLNVKWY--GNFNAVGVNFHVNKDCFGLLGVNGAGKTSTFQMLTGENSISG 1504
QY	1850	EAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEQAQVATAGSG 1909
DB	1505	DAYVNGSVKNNWREAGANTGYCQYDAI IKEMSGEETLYMFARIGIPEKDI PKKNAV 1564
QY	1910	LARLGLSWADRPAGTVSGGNKRLKATALALVGPVAVVFLDEPTTGMDPSARRFLWNSLL 1969
DB	1565	IHAIGIGMYASRQITKYSNGNKRRLSLGIAIVGLPDLVLLDEPTSGVDPKARRIIN 1624
QY	1970	AVVREGSVMLTSHMECEALCSRLATMNGRFRCLGSPQHLKGRPAAGHTLTURVPA 2029
DB	1625	RLRDLGTALVLTSHMECEALCTELAIWYKFCRYGSCQHI KSRYSGSGTLLIRLKNR 1684
QY	2030	R-SQPAAFVAAEPFGSELREAHGRRLFQLPPGRCALARVFGELAVHGAHEGVDFSV 2088
DB	1685	NDAEKTKSTIKQTFRGSVKEEHLVQLNFDIPRQD--SWSRLFKEKLETVSTSLNWDVSL 1743
QY	2089	SQTMLEEVFLYFSKQKQKDEDETEQKAGVVDVDPAGLQH 2128
DB	1744	SQTTLEQVIFBFSRDAGVSSDSEPDGASSTGSANSRQKN 1783

```
RESULT 7
A84845
probable ABC transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84845
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
eusa, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84845
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1816 <STO>
A;Cross-references: UNIPROT:Q9SDB1; GB:AB002093; NID:g6598351; PIDN:AAC02761.2; GSPDB:GN
C;Genetics:
A;Gene: At2g41700
A;Map position: 2

Query Match 16.3%; Score 1813; DB 2; Length 1816;
Best Local Similarity 27.0%; Pred. No. 4.3e-105;
Matches 579; Conservative 287; Mismatches 633; Indels 646; Gaps 70;

QY 260 DSSLSPACSEL-----ICALSHPLSLRLMLRRLKPLILGK---LLFADPTPTTKLMA 309
Db 12 DTTIHPAHSNIDKDTVVEVGKGNPSPEVL-----KLLAEGDFLAFADPTDTNNMID 66
QY 310 QVNRTEELTLDRVREVMELGPRITFTFNDSSNAMLQRLQMQDEGRQRPGRDH 369
Db 67 ILSLKPEPLRV-----TKIP-----KOD 85
QY 370 MEALRSPLDCSGSGYMQDAHDVGLVGLTRVTECLSLDKLEAPSEALVSRALQLL 429
Db 86 IE-LETVI-----TSAH-----YGVSEVRNC-SNPKLGA-----114
QY 430 AEHRFAGVVF--LGPE--DSSDPTHTPDGLGPHVRIKIRMDIDVVTNKRIRDFWD 485
Db 115 -----VVFHEQGHPLFDYSIRLNHTW--AFAGFNNVKSIMD-----TN-----150
QY 486 PGP-----AADPLTLRVYGVGVYLODLVERAAVRV-----LSCANPRAGI---527
Db 151 -GPYINDLENGINTPTMQYSFGSLTLQOVVDSFIIFASQONNDLPLSHNSLSALRFE 209
QY 528 -----YLQOMPYPC--YVDVFLRLVLSRLPLFLTLAWIYSVTLTVKAVVREKTR 576
Db 210 LPWTLFSPSVIRVVPFPTREYTDDEFQSIKSVNGLEL-----247
QY 577 LRDTMRAMGLSRVLMGLWFLSCGLPFLLSAALLVLVLKGLDILPYSHPGVWFLFLAFA 636
Db 248 -----VLCVARDLPAAGR 686
QY 637 VATVQSFLLSAFPSRANLAACGLAYFSILYLPY-----VLCVARDLPAAGR 686
Db 265 LSAIMLSFMISTFTFRAKTAVAVGTLTLFLGAFPPYTYVNDSEVSMVL-----K 312
QY 687 VASLLSPVAFGFCESIALLEECEGAQHNVGTRPTADVSLAQVSGLLLLDAALYGL 746
Db 313 VVASLLSPTAFALGSINFADYERAHVGLRWNSNIWRASSGVSVFVCLL--MWLLDSILYCA 370
QY 747 ATWYLEAVCPGQYGIPEFWNPPFRSSYW-----CGRPKPSAPACPT-----PLDP 792
Db 371 LGLYLDKVLPRENGVRVPMNPIFSKYFGRKKNLQNRIPGPTDMFPADIEVNGQEPDP 430
QY 793 ---KVIIEAPPGLSPG-VSVRSLEKPPGSPQP--ALRGLSLDFYQGHITAFGLHNGAG 846
Db 431 VFESISLEMRQOQLDGRCIQVRNLHKVYASRRGNCCAVNSQLTLTYENQLISLHNGAG 490
QY 847 KTTTSLISGLFPSPGSAFTLGHVDRSSMAAIRPHLGVCPQYNVLFMDLTVDBHWFYG 906
Db 491 KSTTISMLVGLLPTSGDALILGNSIITNMDEIRKELGVCPQHDILPELTVREHLEWFA 550
```

```
QY 907 RLKLSAAVVGPEQDRLLQDVLGVSQSVQTRHLSCGMORKLSVAIAFVGGSSOWILDEP 966
Db 551 VLKGVBEGLSKSTVDMABEVLGDKINTLVRALSGMKRKLGLGIALGNKSVILDEP 610
QY 967 TAGVDVPASRRGIWELLKVKREGRITLILSTHLDDEALLGDRVAVVAGGRGCCGSPFLR 1026
Db 611 TSGMDPYSNRLTWQLIKKIKKRIILLTTHSMDAEELGDRIGIMANGSLKCCGSSIFLK 670
QY 1027 RHUGSYIYTLVKARLPLTNEKADTDMEGSVDTROEKNKGSGSRVGTPLLALVQHWV 1086
Db 671 HHYGVGYTLTLVK-----TSPTVSAAH-----IVRHI 699
QY 1087 PGARLYEELPHELVLVLPYTGADHGSFATLRELDTRLAEL-----RLTGYGSDTSLEE 1141
Db 700 PSATCVSEVGNELISFKLPL--ASLPCFENNFRIESCMKNSDDYFGIYSGISVITLEE 757
QY 1142 IFLKV-----VEECAADTDMEDSGCOHLCTGIAGLDVTLRLKMPPOETALENGEPAG- 1194
Db 758 VFLRVAGCNLDIEDKQEDIFVSPDTKSLVCIG-----SNQSSMQPKLLASCN-DGAGV 811
QY 1195 ---SAPE-----TDQSGPDVAVRGVQHWALTROQLQALLKFLAARRS 1235
Db 812 IITSVAKAFRLIVAAVMTLIGFISIQCCGCSIIISRMFW---RHCKALFIKARSACRD 867
QY 1236 RRGCLPAQIVLPALFVGLALVFSLLIVPPFGHYPALRLSPMYGAQVSEFSEDAQDPGRAR 1295
Db 868 RKTAVQFIIIPAVFLFLGLLF-LQLAPHDPQKSTITUT-----TAYFNLLSGKGGG- 918
QY 1296 LLEALLQEAGLEBPVQVHSHRSFAPEVAEVAKLASGNWTPESPSPACQSQPGARRL 1355
Db 919 -----PIPFD---LSVP-IAKEVAQYI-EGGWI--QPLRNTSYKFPNKEA 957
QY 1356 LPQCPAAAAG---GPPPPQAVTSGEVVQNTGNRLSDFLVKTYPRLV---ROGLKTKKWV 1409
Db 958 LADAIDAAGTLGP-----TLLSMSEFLMSSFDQSQSSREGLSHSDSC 1001
QY 1410 NEVRYGGFSLGGDPGLPQSELGRSVEELWALLSLPLPGCALD-RVLKNLTAWAHSUDAQ 1468
Db 1002 NH-----PDGSLGYTVLHNGTC-----1018
QY 1469 DSLKIWFNNKGMHSMVAFVNRASNAILRAHLPQPARHAHSITTLNHLNLTKEQ-LSEA 1527
Db 1019 -----QHAGPIYINVMHAAILRL---ATGNKNTIQTNRNHLPTKTQRIQR 1063
QY 1528 ALMASSVDVLVSI-CVVFAMSFPASFTVLIERVTRAKHLQLMGGLSPTLYLGNFLWD 1587
Db 1064 DLDAFAAAIIVNI---AFSFI PASFAVPIKEREVKAKHQQLISGVSVLSYLSLTVYWD 1119
QY 1588 MCNLYPACIVLILFAFQQRVAVPANLPALLLLLLYCGWSITPLMYSPFSPVSTPA 1647
Db 1120 FISLFLPSTPAIILFVAFGLEQFIGIGRFLPTVMLELYGLAIASSYTCITFFTEHMA 1179
QY 1648 ---YVLTICNLFIGINGSMATFVLELFDOKLQEVSR1-----1683
Db 1180 QATSSYSVLLPISLVPFSSNVLVHVFSGILMVLVIFSWGLIPATASANSYLKELL 1239
QY 1684 ---LKQVFLIPHFCLGRGLIDMVRNQAMADAPERLG---DRQPSQPLRVEVVGKN----1733
Db 1240 FRYALQNFFRLSPGFCFSDGLASL-----ALLRQGMKDKSGSHGVFEWNVVTGAS1CYL 1291
QY 1734 -----LLAMWIOG---PLFLFTLLLQHSOLLQPPVRSL-----PLLG--1770
Db 1292 GLEVRLEYCRYSMLLISFFHIGIDTKLSLIYIGASRLTELI-YDRVYSTSFSFEPLKDS 1350
QY 1771 -----EBEDVARERERVQATQGDVLVLRNLTKVYRGQR---MPAVDRLCIGIP 1818
Db 1351 TGAISTDMEDDIDVQEDRDVISGLSDNTMLYLQNLKRVYPGDKHGHKPVAVQSLTFSVQ 1410
QY 1819 PGSCFGLLVNGAGAKTSTFRMTVGTDLASGEAVLAGHVSAREPSAAHLSMGYCPQSDAI 1878
Db 1411 AGEFCFGLTNGAGKTTLSMLSGEETPTSGTAFIPGKDIVASPKAIRQHIQYCPQFDAL 1470
```

```
QY 1879 FELTGREHLELLARLGRVPEAQVATAGSGIARLGLSWYADRPACTYSGGNKRKLATAL 1938
Db 1471 FEYLTAKHELELYARIKGVVDHRIDNVVTEKLVFEDLLKHSHKPSFTLGGNKRKLSVAI 1530
QY 1939 ALVGPAVVFLEDEPTGMDPSARRFLNLSLAV-VREG-SVMTLSHSMEECEALCSRLA 1996
Db 1531 AMIGPPVIVLDEPSTGMDFAKRFMDVDVLSRLSTRGKTAVILTTHSMNEAQAALCTRIG 1590
QY 1997 IMVNGRFRCLGSPQHLKGRFAAGHTLTLRVP-----AA 2029
Db 1591 INVGGRLRCIGSPQHLKTRY--GNHLELVFPYNGVKNEVSNVLENFPCQIIQQWLFNV 1648
QY 2030 RSQPA-----2035
Db 1649 PTQPRSLLDLEVCIGVSDSITPDTASASEISLSPEMVQRIAKFLCNEQVRSTLVPLPE 1708
QY 2036 -----AFVAAEPGSELREAHGRLRF 2057
Db 1709 EDVRFDDQLSEQLFRDGGIPLPIFAEWMLTKFKFSAIDSFIOSSPFGATPKSCNGLSIKY 1768
QY 2058 QLPPG-GRCALARVFEGLAVHGAHEGVDFSVSQTMLBEVFLYFS 2101
Db 1769 QLPPGEGSLADAFGHLENNRNLGIAEYSISQSTLETIFNHFA 1813

RESULT 8
C88925
protein F33E11.4 (imported) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C88925
R:anonymouse, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; PMID:99069613; PMID:3851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C88925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1317 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC17542.1; PID:g3158495; GSPDB:GN00023; CESP:F33E11.
C:Genetics:
A:Gene: F33E11.4
A:Map position: 5

Query Match 14.4%; Score 1603; DB 2; Length 1317;
Best Local Similarity 28.3%; Pred. No. 4.le-92;
Matches 447; Conservative 246; Mismatches 511; Indels 378; Gaps 41;

QY 639 TWTSFLLSAPFSRANLAAACGGLAYFSLYLPYLCVAVRDRLPAGGRVAASLLSPVAFG 698
Db 3 SATSATLSVVFW-----LUF--WTAFFSSIDQTPYPLGYRLINCINPDIALN 51
QY 699 FGCESLALLEQEGEAOHNVGTRPAD---VFLSAQVSGILLDAALYGLATVLEAVC 755
Db 52 YGLQLLAAYEQADGLKNGELFTPPSPDNLTFGHALTA--LIVDGIITWIIWIEAVI 109
QY 756 PQQYIPE-PNVPFRPRYWCGRPKSPACPTPLD-----PKVLVEAPGLSPG 806
Db 110 PGEGVPPQKPFVFL-PSYWF-----PNSGSKTVSDSDQFQIQYADHVKLEKEPTDLIPT 164
QY 807 VSVSRLEKRFPGS-----PPALRGLSLDFYOGHITAFIGHNGAGKTTLS 852
Db 165 INVNLTITGTSPPKFLFDCKFGSGGKRAVSNLNKNYPPQCVTLIGHNGAGKSTTFS 224
QY 853 ILSGLFPSPSGSAFTLGHVDRSSMAAIRPHLGVCQYQNVLFMDLTVDVSHVWFYGLKGLS 912
Db 225 MLTGVASPSSGSAVYVDFDITSLPKIRREGLCPQYNTLFGFMTVMEHLEFPKAKL--- 281
QY 913 AAVGPEODR-LLQDVGLVSKQSVQTRHLSGMQKRLSVAIAFVGSGSQVILDEPTAGVD 971
Db 282 ERTWDPPEARILARLRIDFKADFAGALSGQKRLSLAIALIGGSEVVMVLDPTSGMD 341
```

```
QY 972 PASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGRLCCCGSPFLRRHLGS 1031
Db 342 PGAREHTWTLQREKERTILLTHFMEADLLGRIALMAHQLECCCGSPFLKQOYGD 401
QY 1032 GYLLTLVKARLPLTTNEKADTMEGSDVTROBKNGSQGSRVGTQQLLALVQHWHPGARL 1091
Db 402 GYHLTIV-----YDTTSTPDVSKTID-----IIREYIPEAHV 433
QY 1032 VEELPHELVLVLPYTGADGSPATLFRDLTSLAELRLTGYISDTSLEELFLKVVECA 1151
Db 434 FSYIQEATYLLSAT--HRPIFPKLFKELEDHQTCCGITSFGVSITTBEBVLKVGHGTAD 491
QY 1152 ADTDMEDGSCGHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGPD--AVG 1209
Db 492 ERYNVEHG-----IEN--DISEMIEKDDPILODLRAQV 522
QY 1210 RVQGWALTROQLQALLKRLARRRRRLFAQIVLPALFVGLALVFSLIIVPPFGHPAL 1269
Db 523 RVTGFTLQWQAKAMFYKRAIFFFRKWTQFLPQLVFPVAYLVLMVFTSQVLP----- 574
QY 1270 RLSPTMYGAQVSPFSEDAPGDPGRARILLEALLQEALEBPVQSHSRPSAPEVPAEVAK 1329
Db 575 -----SVKE----- 578
QY 1330 VLASGNWTPSPSPACQCSQPGARRLLPDCPAAAGPPPPQAVTGSGEVWQNLTG--RNLS 1388
Db 579 -----QDQTISLAPFSDTKKAG-----HLVDSGNVYVTLTGGSQNLUS 616
QY 1389 DFLVKTYPRLVROGLTKTKWNEVRYGFSLGGDRDPLGSGQELGRSVEELWALLSPLFG 1448
Db 617 SMV-----QGTVTQLGVTVQ-----TVVDITSNVEK----- 641
QY 1449 GALDRVLKNTAWAHSLLDAQDSLKFWNNKGWHSVAFVNRASNAILRAHLPPGPARHAH 1508
Db 642 -----FIMDQTNAM-----GS 652
QY 1509 SITTLNHLNLTKEQLSEAAALMASSVDVLVSIQVVFAMSFVPASPFTLVLEIBRVTRAKHL 1568
Db 653 RTFGLHYALGFVPSMNFSTVSPLKISPAVCV-----AGYSQFLITERKKSKKM 704
QY 1569 QLMGSLSTLYLGNFLNMDCNLYLPACIVLVILFAFQORAYVAPANLALLL--LLLYG 1627
Db 705 QLLSGIRPMPFWLTAFIWDAAWFVIRILCFDAIFVIFNITAYTHDFGVMLILTLFLLYG 764
QY 1628 WSITPLMPASFFSPSPSTAYVVLTCINLFGINGSMAITFVLE----- 1670
Db 765 WTALPFTTYWQFFPESAPKGFMMVTMYHILTMIGSIAVPIISQTSLSLDAGYLWSIIFAM 824
QY 1671 LFSQDKLOEVSRI-----LQOVFLIPPHF-----CLGRGLIDMVNRQAMADAPER 1715
Db 825 LPPTYNISQIATVTFQENVRICKLDCTIPMFKAVTACCG-----TASER 871
QY 1716 L-----GDRQFQSPLRWEVWGNLWAMI-----QGFLFLLTLLQHR--SOLLPOP 1761
Db 872 LYVDNVLFPVGNR-----KGLVYVFLAVQGFYIWMVFMRENDQFTKLFI 919
QY 1762 RVRSI--PLLGEDEEDVARERE-----RVVQATQGD--VLVRLNLTKVTRGQRM 1807
Db 920 RCRKADNPIMDITDVKYDERDVEDSDVIAEKSVVQRLANNKNTALVSNLWKVY--GNF 977
QY 1808 PAVDRLCIGIPPGSCFGLGVNGAGKTTSTFRMTVGTDLASGEAVLAGHSVAREPSAHL 1867
Db 978 NAVGVNPHVNSKCFGLLVNGAGKTTSTFQMLTGENSEISSGDATVNGVSNVKNWREQA 1037
QY 1868 SMGYCPSQDAIFELLTGREHLELLARLGVPEAQVATAGSGIARLGLSWYADRPACTYSG 1927
Db 1038 NTGCPQYDAIIKEMSGEETLYMFARINGIPEKIDIPKKVNAVIHAIGIGVMSRQIKTYS 1097
QY 1928 GGNKRLKATALVGDPAVFLDEPTGMDPSARRFLNLSLAVVREGSRVMTLSHSMEE 1987
Db 1098 GGNKRRRLSLGIAIVGLPDVLLDDEPTSGVDPKARRIINILNRLDLGTALVLTSHSMDE 1157
```


Qy 1988 CEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAR-SOPAAAFVAEPGSGE 2046
Db 1158 CEALCTELAIWVYKFRGFCYGCQHIKSRYSYGSYTLIRLNRNDAEAKTKSTIKQTFRGV 1217
Qy 2047 LREAHGRLTFQLPFGRCALARVFGELAVHGAHGVDEDFSVQSTMLEEVFLYFSKQDGK 2106
Db 1218 IKBEHVLQMLNFDIPROGD-SWSKLFKEKLTVSTSLNWDYSLQTTLEQVIFESRDAGV 1276
Qy 2107 DEDTEQKEAGVGVDPAAGLQH 2128
Db 1277 SSDSEPDGASSTGSANSRGKQN 1298

RESULT 9
T15200
hypothetical protein F12B6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
C;Accession: T15200
R;Pauley, A.; Maggi, L.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid F12B6.
A;Reference number: Z18307
A;Accession: T15200
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1447 <PAU>
A;Cross-references: EMBL:AF003138; NID:G2088708; PID:G2088709; PIDN: AAB54153.1; GSPDB:GN
A;Experimental source: strain Bristol N2; clone F12B6
C;Genetics:
A;Gene: CESP: F12B6.1
A;Map position: 1
A;Introns: 79/2; 114/3; 177/1; 224/3; 331/1; 345/3; 373/2; 417/2; 464/1; 536/1; 659/2; 6
C;Superfamily: ATP-binding cassette homology

Query Match 14.4%; Score 1600.5; DB 2; Length 1447;
Best Local Similarity 27.3%; Pred. No. 6.8e-92;
Matches 475; Conservative 257; Mismatches 533; Indels 477; Gaps 51;

Qy 492 PLTDLRVWGGFVYLO-----DLVERAAVRVLGSG-ANPRAGLYLOQMPY 534
Db 13 PLIDLKVITGFSPLQSGWFSLOKSKSLSDISIDRAIMSELTNTDANLGVIAQOEY 72
Qy 535 PCYVDDVFLRVLSPLFLTLAMIYVTLTKAVVREKETRLDTRMAGLSRAVLWLG 594
Db 73 PCTVKDTFNVAL--FMELFLLISFIPSAALLVKNIVYEKQIKEQWRAMGLGDVAHFIS 130
Qy 595 WFLSCLGPFLLSAALLVLVLKGLDILPYSHGVVFLFAAFVATVTSQFLLSAFTSRAN 654
Db 131 WGLISLVNFTISVLIIISIKVAKIFDYDTYTLFLVLILFSSIAMSIFFSTLTFTNAN 190
Qy 655 LAAACGGALFYLSLVPYLVCMVREDRLPAGGRVAASLLSPVAFGFCESLALLEQGEA 714
Db 191 IATAATCVLWFVFFIPQLLRTDRISPTFNRLS-ILPPTAMGH-CFKULESFNAMERA 248
Qy 715 QWNVG--TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPGQYIGPEPNWFFRRS 772
Db 249 TWSDLWENNPVLGI-SVELCMIMLVVDVAVFLILAWYISAVAFGDEGVQPLWFFTLK 307
Qy 773 YWCGRPKPKPAPCPTPLDPKVLVEAPGLSGVGSVRSLE----KRPFGSP-----Q 821
Db 308 YW-----APGLYKNRVEFVDDHFDFTIPNSDGSFDS 338
Qy 822 P-----AIRGLSLDFYOGHITAFILGHNGAGKTTTLTSLGFPSPGSGSAILGHDVRRSM 876
Db 339 PTNLTLALDCLNLRLEYEQITGLLGHNGAGKTTTSLCGLYAPSSGTAKIYORDIRIDL 398
Qy 877 AAIRPHLGVCQYNNVLFMLTVDBHWYFYGLKLSAAVVGPEQDRLLQDVGLVSKQSVQ 936
Db 399 RRVRDVLGICPQHNVLFSHLTVSQELRLFAALAGVPDSBELTSQVDEILASVSLTERANKL 458
Qy 937 TRHLSSGQMRKLSVAIAFVGGQSVVILDEPTAGVDPASRRGIWELLKUYEGRTLLILSTH 996

Db 1443 -GRFTAVNELCLAVDQKECGLLGVNGAGCTTTTNILTGOSFASSGEAMIGGRDVTE--- 1498
Qy 1864 AAhLSMGYCQSDAIPELLTGREHLELLALRLGVPEAOVAQTAGSLARLGLSWYADRP 1923
Db 1499 --LISIGYCPQFADMLDLTGRESLEITLAQMGFENYKA--KAEILILECVGMIAHADKLV 1554
Qy 1924 GTYSGGNKRKLATALALVGPVAVFDELPTTGMDPSARRFLNLSLLAVVREGS-VMLTS 1982
Db 1555 RFYSGGGRKISVGVALLARTQMIIIDEPAGIDPKARREVWELLWCRESHSALMLTS 1614
Qy 1983 HSMEECEALCSRLAIVMNGFRCLGSPQHLKGRFAAGHTILTLRV--PAARSQPAFAAFAVA 2040
Db 1615 HSMDECEALCSRATLVNRSGLIALGSSQELKSLYGNNTWTLISLYEPNQDM-VVOLVOT 1673
Qy 2041 EFPGESELEAHGGR---LRFQLPPGRCALARVFGELAVHGAHEGVDFSVQTMLEEVF 2097
Db 1674 RLPNSVLKTTSTNKLNLKQIPEKEDCNSAKFEMVQALAKDLGVKDFILAQSLEETF 1733
Qy 2098 LYFSKDOGQKDED 2109
Db 1734 LRLA---GLDED 1742

RESULT 11
S60124
transport protein homolog C48B4.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 13-Jan-1996 #sequence revision 12-Apr-1996 #text_change 16-Aug-2004
C/Accession: S60124; S40724; S40725
R/Kershaw, J.
submitted to the EMBL Data Library, November 1995
A/Reference number: S60124
A/Accession: S60124
A/Molecule type: DNA
A/Residues: 1-1767 <KER>
A/Cross-references: EMBL:Z29117; NID:g439247; PID:g1066912
C/Genetics:
A/Map position: III
A/Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/
C/Suprafamily: ATP-binding cassette homology
C/Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein
F:628-818/Domain: ATP-binding cassette homology <ABC1>
F:645-652/Region: nucleotide-binding motif A (P-loop)
F:764-769/Region: nucleotide-binding motif B
F:1457-1642/Domain: ATP-binding cassette homology <ABC2>
F:1474-1481/Region: nucleotide-binding motif A (P-loop)
F:1586-1591/Region: nucleotide-binding motif B

Query Match 12.68; Score 1407; DB 2; Length 1767;
Best Local Similarity 25.98; Pred. No. 1.3e-79;
Matches 485; Conservative 298; Mismatches 665; Indels 424; Gaps 59;

Qy 383 GYSQDAHA--DVGHVLTGLRVTECLSLD---KLEAAPGEAALVSR-ALQLLAE----- 431
Db 159 GYTSKDAARAKTVDLKKFAERQSAKLKLVNSESQELLTLVRLNDLPLMNETFCAI 218
Qy 432 HRFWAGVVF-----LG---PESSDPTHEPTDLPQGHVRKIRMDIDVVT 474
Db 219 NSVAAGVVDFVDVTKNKLNYRILLGKTPBETHLTSTYNYGPSSGR----- 267
Qy 475 RTNKIRDFWDPPGAADPLTLRVVWGGFVYLODLVERAAVRVLSGANPRAGLVLOOMPY 534
Db 268 -----YSRIPSSPP-----YWTSAFLTFQHAIESFLSSVQSGAPDLPITRLGPE 313
Qy 535 PCYVDDVFLRSLRSLFLTLAMIYSVTLTKAVVRE---KETRLRDTWRAMGLSRAVL 591
Db 314 PRYKTSVSFAID-----FPPFIWAFVTFINVIHITREIAENHAVKPYLTANGLSTFMF 368
Qy 592 WLGHFLSLGFLLSAALLVLVLKGLDILPVSHPGVVFLLFAFAVATVTSQFLLSAPFS 651
Db 369 YAAHVMAFLKFFV--IFLCSIIPLTFMFEVSPAAIIVTLMYGLGAVIFGAFVASFFN 426
Qy 652 RANLAAACGGGLAYFSLVLPYVLCVAVRDLR-----PAGGRVAASLLSPV-----AFGFG 700

Db 427 NTNSAIK-----AILVANGAMIGISYKLRPELDQISSCFLYGLNINGAFALA 473
Qy 701 CESLA-----LLEQEGEQAQHNHVTTRTADVFLSAQVS-GLLLID-----AALYG 745
Db 474 VEAISYMRREERELNTNMFNDSLSLHFSGLWALVMMIVDILMMSIGALVVDHIRTSDAFS 533
Qy 746 LATWY-----LEAVCPQOYGIPEPWNPFPRSYWCGPRPKSPAPCTPLDPKV 794
Db 534 LRTLDFEAPEDDENOTDGVTAQNTINEQVRNRVRSD-MENMPMASTSLNPNADS 592
Qy 795 LVE-----EAPGLSPGVSRSLERKFRPGSPALRGLSLDFYQGHITAFGLHNGAGKT 848
Db 593 LLSGSTEADGARTARADIIVRNVLKWTSTTGERAVDGLSLRAVRGQCSILLGHNGAGKS 652
Qy 849 TTLSIISGLFPPPGGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFMDLTVDHWFVGRLL 908
Db 653 TTFSSIAGIIRPTNGRITTCGYDVGNPEGETRHHGMCPOYNPLYDQLTVSEHLKLYVGL 712
Qy 909 KGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGMOKLSVAIAFVGGSOVILDEPTA 968
Db 713 KGAREKDFKODMKRLLSDVKLDFKENEKAVNLSSGMRKRLCVCMALIGDSEVLLDPTA 772
Qy 969 GVDPASRRGIWELLKYREGRTILSLTHLDEAELLDGDRVAVVAGGRLCCGSPFLFRRH 1028
Db 773 GMDPGARQDVQKLVREKANKRTILLTHYMDAERLGDWVFINSHGKLVASGTINQYLKQK 832
Qy 1029 LGSGYLTLVKARLPLTTNEKADTMEGSVDTQEKKNQSGQSRVGTPTQLLALV-----Q 1083
Db 833 FGTGYLLTVV-----LDHNGDKRK-----MAVILTDVCT 861
Qy 1084 HWTPGARLVPELPHLVLPYTGADHSGPATIFRELD-----TR 1123
Db 862 HYKAEARGEMHQOIEIILP-EARKKEFVLPFOALEAQDRNRSNVFDMNPTLKSQ 919
Qy 1124 LAELRLTGYGISDTSLEEIFLKVVECAADTMEDGSCGQHLCTGTAGLDTVLRLKMPQ 1183
Db 920 LATLEKRSFGLSLNLTQEQVITI-----GDKVDKAIASRQ-NSRISHNSR 963
Qy 1184 ETALENGEPAG-----SAPETD-----QSGPDPAVGRVQGWALTRQQLALLKRF 1230
Db 964 NASEPSLKPAYDTQOSTKTSADSYQKLMDSQARGEK-----SGVAKMVAQFISIMRKKFL 1019
Qy 1231 LARRSRGLFAQIVLPALFVGLALVFSLIIVPPGHYPALRLSPITMYCAQVSFFSEADAPG 1290
Db 1020 YSRNNAQLFTQVLIPIILGLV-----GSLTTLKSNNT-----D 1054
Qy 1291 PGRARLLEALLQEALEEPVQHSRRFSAPVPAEVAKVL-ASG-----NWTPEPSPA 1344
Db 1055 QFRS-----LTPSGIEPSKVVRWFENGITPEEAANFEKILRKSGGFEVLNNTKNP--- 1105
Qy 1345 CQCSQPCARLLPDCPAAGPPPPQAVTGSGEVQVQLTGNLSDFLVKTYPRVLROGLK 1404
Db 1106 -----LPNITKSLIGEMPPTI----- 1122
Qy 1405 TKKVVNRYVYGGSLGDRDPLPSGQBLGRSVEELMALLSPLPGALDRLVKNLTAWAHS 1464
Db 1123 -----GMTMN----- 1127
Qy 1465 LDAQDSLKIWFNKNKGWH---SMVAVNVRAS-NAILRAHLPPGPARHAHSITTLNHPNLT 1520
Db 1128 ---SDNLEALFNRYHYHVLFTLSMINARLGTVDAEISSGVFLYSKTSN-----SNLL 1180
Qy 1521 KEQISEAALMASSVDVLVSVTCVVFAMSFPASPFTVLIERVTRAKHLQLMGGLSPFLYW 1580
Db 1181 PSQI-----IDVLLAPMLLIFAMVTSFTFMFLIEERTCQFAHQOQLTGISPTIFY 1231
Qy 1581 LGNFLNDCMNYLPACIVLILFIAFOORA VAPANLPALLLLLLYCWSTPLMYPASFF 1640
Db 1232 SASITYDGIIYSLCLIFLFWLAF-----HMYDHLAIVILFWFLYFSSVPFIYASFL 1287
Qy 1641 FSVFSTAYVVLTCINFIGINGSMAITFVLELFDQKLOE-VSRILKOVFL-IPPHFCIGLR 1698

Db 1288 FQSPSKANVLL--IIQVVISGAALLAVFLIFMIFNDEWLKSLIILNIFMFLPSYAFGS 1345
Qy 1699 GLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKILLAMVIOQL-FLIFTL----- 1750
Db 1346 AII-TINTYGMLPSEEL-----MNDHCGKNALWMTGVCVSFALFVLLOPKFYR 1396
Qy 1751 -----IQHRSQILPQPRVRSPLPLGDEDEDVARERERVQATGQDVLVLRNLTKVYR 1803
Db 1397 FLSQVMTVRSSHHNVQPMGDLPCV-----ESVSEERERVHRVNSQNSALVIKDLTKTF- 1451
Qy 1804 GORMPAVDRLCLGIPGECFGLLVNGACKTSTFRMVTGDTLASGEAVLAGSHVAREPS 1863
Db 1452 -CRFTAVNELCLAVDOKECFGLLVNGACKTTFNLTGQSFASGEAMIGRDVTE--- 1507
Qy 1864 AAHLGMYCQPSDAIFELLTGREHLELLARLRGVPEAQVAQTAGSGLARLGLSWADRP 1923
Db 1508 --LISIGYCQFDALMLDITGRESLEILAQMHEFNKYA--KAEILECVGMIAHADKLV 1563
Qy 1924 GTYSGNKKKLATALVGDPAVFLDEPTTGMDSARRFLNWSLLAVVBGRS-VMLTS 1982
Db 1564 RFYSGGQKRIKISGVALLAPTQMIILDEPTAGIDPKARREVWELLIMCKREHSNLSALMLTS 1623
Qy 1983 HSMERCEALCSRLATMVGRCFLGSPQHLKGRFAAGHTLTV--PAARSQPAAFVAA 2040
Db 1624 HSMDECEALCSRIAVLNRLSLTAIGSSQELKSLYNNYTMTLGSLYEPNQORDM-VVOLVOT 1682
Qy 2041 EPPGSELREAHGGR---LRFQLPQRCALARVFGELAVHGAHGVDEPVSQTMLEEVF 2097
Db 1683 RLPNSVLKTTSTNKTNLNKWQIPKEKEDCWSAKFEMVQALAKDLGVKOFILAQSSLETF 1742
Qy 2098 LYSKDKQKQD 2109
Db 1743 LRLA---GLDED 1751

RESULT 12
T42749
ATP-binding cassette transport protein homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2004
C:Accession: T42749
R:Wu, Y.C.; Horvitz, H.R.
Cell 93, 951-960, 1998
A:Title: The C. elegans cell corpse engulfment gene ced-7 encodes a protein similar to A
A:Reference number: Z22259; MUID:98297348; PMID:9635425
A:Accession: T42749
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1704 <WUY>
A:Cross-references: UNIPROT:P34358; EMBL:AF049142; NID:g3172340; PIDN:AAC24116.1; PID:g3
A:Note: ced-7
C:Genetics:
C:Superfamily: ATP-binding cassette homology

Query Match 12.6%; Score 1405; DB 2; Length 1704;
Best Local Similarity 25.9%; Pred. No. 1.6e-79;
Matches 486; Conservative 297; Mismatches 668; Indels 422; Gaps 59;

Qy 383 GYSQDAHA--DVGHVLVTGLRVTECLSD---KLEAAPSEALVSR-ALQILAE----- 431
Db 92 GYTSKDAAKRTVDLDMKKFAERFOSAKLISVRNSESBEQLTLVLRNDLPMLNETFC 151
Qy 432 HRFWAGVVF-----LG--PEDSSDPTHEPTDLPDGPVHRIRKIRMDIDVVT 474
Db 152 NSYAGVVVDEVDVNTKKNLVRILLGKTPEETHLTETSYNYPGSSOR----- 200
Qy 475 RTNKIRDFWDGPAADPLTDLRYVWGQVLYQDLIVERAAVRVLSGANPRAGLYLQOMPY 534
Db 201 -----YSKIPSSP-----YWTSAPLTFQHAIESFLSSVQSGAPDLPITLRGLPE 246
Qy 535 PCYVDDVFLVLSRLPLFLTLAWIYSVTLTVKAVRE---KETRLRTMRAAGLSRAVL 591
Db 247 PRYKTSVSAFID-----FPPPIWAFVTFINVIHITREIAAENHAVKPYLTAMGLSTPMF 301

Qy 592 WLGFWSLCLGPFLLSAALILVLVLKGLDILPYGHPGVVFLFLAFAVATVTSFLLSAFFS 651
Db 302 YAAHVWMAFLKEFV--IFLCSIIPITFVMEFVSPALIVTVLMYGLGAVTIFGAFVASFFN 359
Qy 652 RANLAAACGLAYFSLYLPYLCVARDRL-----PAGGRVAASLLSPV-----AFGFG 700
Db 360 NTNSAIK-----AILVAGAMIGISYKLRPELDQISSCFYGLNINGAFALA 406
Qy 701 CESLA-----LLEEOGEGAQWNVGTPTADVFSLAQVS-CLLLLD-----AALYG 745
Db 407 VEASIDYMRERELNLTNMFNDSSLSHFSLGWALVMVINDILMWISGALVVDHIRTSADES 466
Qy 746 LATWT-----LEAVCPGOYIPEPMFPFRS--YWGPRPKPKPAPCPTPLDPK 793
Db 467 LRTLDFEAPEDDENQTDGVTQAQTRINEOVNRVRSRDMETQMPMASTSLNPPNADSD 526
Qy 794 VLVE-----EAPPGLSPGVSVRSLEKFPSPQALRGLSLDFYQGHITAFIGHNGAGK 847
Db 527 SLLEGSTEADGARDTARADIIVRNLVKIWTSTGERAVDGLSLRAVRGQCSILLGHNGAGK 586
Qy 848 TTTLISLGLFPSPGSAFILGHVRSWMAAIRPHLGVCPQYNVLFDMLTVDHWHVWFYGR 907
Db 587 STTFSSIAGIIRPTNGRITICGYDVGNBEGETRRHIGMCPQYNPLYDQLTVSEHLKLVYG 646
Qy 908 LKGLSAAVVGPEQDRLLOQDVLVSKQSVQTRHLSGMOKKLSVAIAFVGGSVVILDEPT 967
Db 647 LKGAREKDFQDMKELLSDVKLDFXENKAVNLSGMRKLCVCHALIGDSEVVLDEPT 706
Qy 968 AGVDPASRRGIWELLKYREGRTLILSTHLDDEALLGDRVAVAGGRGCCGSPFLRR 1027
Db 707 AGMDPGARDQVKLVEREKANRTILLTHYNDSEARLGDWVFMHSGKLVASGTNYLQK 766
Qy 1028 HLGSGVYLTVKARLPLTNEKADTDMESGVDTRQEKNGSQGSRVGTPLLALV----- 1082
Db 767 KFGTYLLTV-----LDHNGDKRK-----MAVILTDVC 795
Qy 1083 OHWVPGARLVBELPHELVLVLPYTGAHDGSAFATLFELD-----T 1122
Db 796 THYVKEARGEHWGQOIEIILP--EARKKEFVPLFQALEAQDRNYSRVNFMNPTLKS 853
Qy 1123 RLAEIRLTGYGISDTSLEBIFLKVVECAADTDMEDSGQHLCGTGAGLDVTLRKMP 1182
Db 854 QLATLEMRSGFLSNTLSQVFTI-----GDKVDKAIASRQ--NSRISHNS 897
Qy 1183 QETALENGEPAG-----SAPETD-----QGSQPDVAVRGVQHWALTRQQLALLKRF 1229
Db 898 RNASEPSLKPGAYDTQSTKSADSYQKLMDSQARGPEK-----SGVAKMVAQFISIMRKF 953
Qy 1230 LLARRRRRLGPAQIVLPALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVFFSEADAPG 1289
Db 954 LYSRRNWAQLTQVLPIILGLV-----GSLTTLKSNNT---DQFSVRS----- 995
Qy 1290 DPGARLLLEALLOAGLEBPPVQHSRHSFAPEVPAEVAKVL-ASG-----NWTPESPSP 1343
Db 996 -----LTPSGIEPSKVWVRFENGTTIPEEAAANFEKILRKGGFEVLNNTKNP-- 1042
Qy 1344 ACQCSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQNLGTGNLSDFLVKTYPLVRQGL 1403
Db 1043 -----LPNITKSLGEMPPATI----- 1059
Qy 1404 KTKKWNVEVRYGGSLSGGRDPGLPSGBQLGRSVEELWALLSPLPGGALDRVLKNTAWAH 1463
Db 1060 -----GMTN----- 1064
Qy 1464 SLDAQDSLKIWNKNGWH---SMVAFVNRAS-NAILRAHLPPGPARHAHSITTLNPLNL 1519
Db 1065 ---SDNLEALFNRYHVLPITLISMINRARLTGTVDABEISSGVFLYSKSTSN---SNL 1116
Qy 1520 TKEQLSEALMASSVDVLVSVICVVFAMSFPVASFVLVLEIBERVTRAKHLQMGSLPLY 1579
Db 1117 LPSQ-----IDVLAEPMLILIPAMVTTSTFVFMLEIBERTCOFAHQQLTGISPTTF 1167


```

QY   2017 AAGHTLTILRVP-----AARSQPAA-----2035
      | : |||||
Db   1041 --GNHLELEVFFYNGVKPKPNSVNELENFCQIIHQWLFPNVPTOPRSLLGLDLEVCIGVSDS 1098
      | : |||||
QY   2036 -----
Db   1099 ITPDTASAEISLSPEMVQRIAKFLGNEQRVSTLVPLPEEDVRFDQLSEQLFRDGQIP 1158
      | : |||||
QY   2036 -----AFVAAPFGSESLREAHGRLRFQLPPG-GRCALARVFGEIAY 2076
      | : |||||
Db   1159 LPFAEWLTKFKPSALDSITQSFFPGATPKSCNGLSIKYQLPFEGGSLDAFAGHLER 1218
      | : |||||
QY   2077 HGAEHGVEDFSVSQTMLEEVFLYES 2101
      | : |||||
Db   1219 NRNRGLGIAEYSISQSTLETIFNHFA 1243
      | : |||||

RESULT 14
T46467
hypothetical protein DKFZp3434E1030.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Aug-2004
C:Accession: T46467
R:Blum, H.; Baurtsachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: P223034
A:Accession: T46467
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-269 >AAA>
A:Cross-references: UNIPROT:Q9NT80; EMBL:AL137481
A:Experimental source: adult testis; clone DKFZp3434E1030
C:Genetics:
C:Superfamily: ATP-binding cassette homology

```

```

Db      923 -----FIVKVNK 929
Qy      1397 R-----LVROGLKTKWVNEVYGGFSLGRDPGLPSGOELGRSVEELWALLSPLPGA 1450
Db      930 KGDskFYKISQGDKNAILMNIITASAMYL--RDPsvTK-----965
Qy      1451 LDRVUKULTAWHSLDAQDSLKWFNN-----KGWHSMAFVFNPNASNAIIRAHLPQPAR 1505
Db      966 ----LPHVT-----SRVIMNDPKIKYGLASFLLFEN-----994
Qy      1506 HAHsITTLNHLNLTKEQSEALMASSVDVLSICVVFAMSFPASFTIIVLIEERVTRA 1565
Db      995 -----IFFLLVLAGIFIQS--TVVLIIEKICKF 1020
Qy      1566 KHLQMGGLSPTLYWLNFLWDMCNLYLVPACIVVLIFLAFQORAYVAPANLPALL--LLL 1623
Db      1021 AHQYLTGLSTIAYGWVFLWDF-----LLFTFELLYTIGFLISFGVLQGHIEIVV 1072
Qy      1624 LLYG--WSITPLMYPASFFSVSTAYVLTICNLFIGINGSMATFVLELFSQKLOEVS 1681
Db      1073 IFYGLLFYFAPLVLTSTALINTPTGNFLL---YMFCCIPWLAYSIVSELHNPFPPIQKYS 1129
Qy      1682 RILKQVFLIF-PHFCLGRGLIDMVR---NOAMADAPERLGRDQFQSPLRWEVVGKULLA 1736
Db      1130 DEIEYGRIFNPSIGFLAGLMKTAALNPKSGLDKHEHILTNLWYEGIFFE-----LM 1183
Qy      1737 MVIQGPLLFLFTLLL-----OHRSQLPQPRVRSPLPLGGEDEEDVARER 1780
Db      1184 FLFFGGIFL--TILGCATLKPFRACFRGTRRRSQ--PREKR--YKGIESCkAVKEE 1236
Qy      1781 ERVVOGATQGD-VLVRLNLTkVRGQRMpAVDRICLIGIPGECFGLLVNGAGKTSTFRM 1839
Db      1237 EQLVQVEDKNETVLVIDGLVKDF--GKFRAVNDLSISVGHCECFGLMANGAGKTTTFDI 1294
Qy      1840 VTGDTLASRGEAVLAGHSVAREPSAHLNMGYCPCQSDAIFELLTGREHLELLARLGVPE 1899
Db      1295 ITGLTMTGGSATIDGHDITET-----IHGYCQFQDAMLQOISCRQTLIRIMAKLOQYP- 1348
Qy      1900 AQVAQTAGSLARLGLSWYADRPAGTYSGGNKRKLATALAVGDPVAVVFLDEPTGMDDPS 1959
Db      1349 -NVKEVVVELVDCVMSDFGKLVKNCSGGQKRKISVGIALMSRATCIILDEPTAGIDPR 1407
Qy      1960 ARPLWNSLLAVVREGR-----SVMLTSHSMECEALCSRLAIVNNGRFRCLGSPQHLKG 2014
Db      1408 ARREIND-----IHEMEQAKCSIVLTSHSMECEALCTRIGILRKGMIALGTSQSLKS 1463
Qy      2015 RFAAGHTLTLRVPAARS-OPAAAFVAAEFPGSELREAHGG---RLRFOLPPGGRCALARV 2070
Db      1464 QYGNTRYMTLILNSLEDLESVCVIVSEEMPDVAVLKTPESSILTSIVWELPKSKSDKWSEK 1523
Qy      2071 FGELAVHGAHGVBDFSVQTMLEVEFLYFSKQDKDEDTEEQKEA 2116
Db      1524 YNQVEVLAKKANAKDYMLTQASLEDTFIRLI-----TTEEEBEA 1562

```

Search completed: January 3, 2005, 10:55:16
 Job time : 69.0254 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:24:44 ; Search time 27.7979 Seconds
(without alignments)
5119.750 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMVRRR.....QHPKRVSQFLDDPSTAETVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5768.5	51.8	2261	US-09-526-193A-1	Sequence 1, Appli
2	5747.5	51.6	2261	US-09-032-438C-118	Sequence 118, App
3	4936	44.3	2273	US-09-032-438C-3	Sequence 3, Appli
4	4858	43.6	2235	US-09-032-438C-6	Sequence 6, Appli
5	3818.5	34.3	1375	US-08-665-259-26	Sequence 26, Appl
6	3818.5	34.3	1375	US-08-762-500-26	Sequence 26, Appl
7	2694.5	24.2	1457	US-08-665-259-27	Sequence 27, Appl
8	2694.5	24.2	1457	US-08-762-500-27	Sequence 27, Appl
9	2693.5	24.2	1472	US-09-032-438C-119	Sequence 119, App
10	2507	22.5	1704	US-09-032-438C-120	Sequence 120, App
11	2502	22.5	1704	US-08-762-500-75	Sequence 75, Appl
12	2501.5	22.4	1684	US-08-665-259-25	Sequence 25, Appl
13	2501.5	22.4	1684	US-08-762-500-25	Sequence 25, Appl
14	430	3.9	162	US-10-000-489-102	Sequence 102, App
15	427	3.8	328	US-09-724-797-8	Sequence 8, Appli
16	414	3.7	607	US-09-252-991A-18351	Sequence 18351, A
17	393.5	3.5	588	US-09-489-039A-13579	Sequence 13579, A
18	389	3.5	589	US-09-328-352-7592	Sequence 7592, Ap
19	375	3.4	594	US-09-543-681A-5528	Sequence 5528, Ap
20	374.5	3.4	308	US-09-602-787A-348	Sequence 348, App
21	371.5	3.3	335	US-09-252-991A-20837	Sequence 20837, A
22	363.5	3.3	392	US-09-602-787A-352	Sequence 352, App
23	357.5	3.2	271	US-09-602-787A-350	Sequence 350, App
24	354.5	3.2	315	US-09-328-352-4388	Sequence 4388, Ap
25	350.5	3.1	788	US-09-252-991A-28171	Sequence 28171, A
26	335	3.0	317	US-09-489-039A-10626	Sequence 10626, A
27	333.5	3.0	929	US-09-252-991A-22946	Sequence 22946, A

ALIGNMENTS

RESULT 1

US-09-526-193A-1
; Sequence 1, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pimstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526,193A
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-526-193A-1

Query Match 51.8%; Score 5768.5; DB 4; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;

Qy	1	MAFWTQLMLLWKNFMVRRRQPVQLLVLLWPLFFFLVAVRSHHPLEHHECHFPNKP	60
Db	1	MACPQRLLLWKNLTFRRQTCQLLLEVAWPLFFLLISVRUSYPPEHCHFPNKA	60
Qy	61	LPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRSLNFNDSLVSRLLADARTVLGASAH	120
Db	61	MPSAGTLPWQGLICNANNPCFRYPTGCEAGVGVGNFNKSVARLFSDARRLLYSQD	120
Qy	121	TLAGLGLKIATLRAARSTAQ-----POPTKQSPLEPML--	154
Db	121	SMKDMRKVLRTLQOIKKSSNKLQDFLVDNETFSGFLYHNLSPKSTVDKMLRADVILH	180
Qy	155	-----DVAEL-----LTSLL	164
Db	181	KVFLQGYQLHLTSLCNKSKSEMITQLGQDQVSELCGLPREKLAARVLRNSMNDLKPI	240
Qy	165	RT--ESLGLALGQAEPLHSLLEAAEDLAQELLALRSILVELR---ALLQRPRTGSGPLEL	219

Db 241 RTLNSTSPFFSKELAEATKILLHSLGTLAQELFMRSSWDMRQWELFNTVNNSSSTQI 300
Qy 220 ---LSEALCSVRGSPSTVSGPSLWYASDLMLVGV---QEPESALPDSSLSPACSELIG 272
Db 301 YQAVSRVICHGHEGGGLKISLWYEDNNYKALFGNGTEDEAETFYDNSTTPYCNDLMX 360
Qy 273 ALDHPILSRLRLWRLKPLIIGLKLPAEDTPTTRKMAQVNRTPFELTLLDRVREVMIG 332
Db 361 NLESPLSRIIWKALKPLLVGLKILYTPDTATROVMAEVNKTFOELAVFVHLEGMWEELS 420
Qy 333 PRIFTEWNSNVAMLQRLQOMQDEGR-RQPRPGRD-----HWEALRSFLDP 379
Db 421 PKIWTFWENSQEMDLVRMLDSRUNDHFEOQLDGLDWTQADIVAFIAKHPEDVQS---- 476
Qy 380 GSGG--XSWQDAHADVGHVLTGLRVTECLSLDKLEAAPSEAAALVSRAQOLLAHREFWAG 437
Db 477 -SNGSVITWREAFNETQARTISRFEVCYNLKNLEPIATEVWLNKSMELLDERKEFWAG 535
Qy 438 VVFLGPEDESDPTBHTPDGLPGHVRJIRKMDIDVTRTKIRDRFWDPGGAADPLTDLR 497
Db 536 IVFTGITFGSIELPH-----HVYKIRKMDIDNVERTNKIKDGYWDPGPRADPPEDMR 587
Qy 498 YVMGFFVYLODLVERAAVRVLGSANPRAGLYLOQMPYPCYVDDVFLRLVLSRLPLFLTLA 557
Db 588 YVMGFFAYLODVVEQAIIRVLTGTTEKKTGYVMQMPYPCYVDDVFLRVMGRSMPLFMILA 647
Qy 558 WIYSVTLTVKAVVREKETRLDRMRAMGLRAVLWGLFSLCGLPFLLSAALVILVILKLG 617
Db 648 WIYSVAVIIGIYVEKEARUKETMRINGLONSILWFSMFISSLIPLLVASGLLVILKLG 707
Qy 618 DIIPYSHPGVFLPLAFAVATVTOQSFLLSAFAFSRANLAACGGLAVFSLYLPVILCVAW 677
Db 708 NLLPYSDFSVVVFVFLVAVVVTILQCEFLISTESRANLAACGIIYFTLYLPVILCVAW 767
Qy 678 RDLRPAAGRAVAASLLSPVAFGFCESJALLBEEQEGAGQWNVGRPT-ADVESLAQVSGL 736
Db 768 QDYVFTLKLIFASLLSPVAFGFCGEYFALPEEQIIGVQWNLNLFESPVVEEDGFNTTSSVM 827
Qy 737 LLLDAALYGLATWYLEAVCQGYGIPPEWMPFRSRYWCGPRPPKSPACPTPLD-PKVL 795
Db 828 MLFDTFLYGVMTWYIEAVFPQYIGIPRPWYFPCTKSYWFGESDEKSHPGSNQKRISEIC 887
Qy 796 VEAPPGLSGVSVRSLEKFPSPQALRGLSLDFYQCHITAFPLHNGAGKTTLSILS 855
Db 888 MEEPTHLKLGVSITQNLVRYRDMKVAVDGLALNFYEQITSPFLHNGAGKTTTWSILT 947
Qy 856 GLFPSPGGSFILGHVRSMAAIRPHLGVCPQVNVLFMDLTVDEHVWYFGRKLGLSAAV 915
Db 948 GLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFMDLTVVEHIWFWYARLKLSEKH 1007
Qy 916 VGPEQDRLLQDVLG-VSKQSVOTRHLSSGMQRKLSVAIAFVGGSQVVLDEPTAGVDPAS 974
Db 1008 VKAEMEQLMDVGLPSSLKSKTSQLSGGMQRKLSVALAFVGGSKVVLDEPTAGVDPS 1067
Qy 975 RRGIIWELLKYRGRTILSTHLDDEAILGDRVAVVAGRLCCCGSPFLRHRHLSGYY 1034
Db 1068 RRGIIWELLKYRGRTIILSTHMDADVLGDIRIAIISHGLCCVGSLSFLKQNLGTGY 1127
Qy 1035 LTLVKARLPLTTN-----EKADTDMGSDVTRQEKNGSGSVRGTPQLLALV 1082
Db 1128 LTLVKDVESSLSCRNSTSVLKKEDSVSQSSSDAGLSGHESDTLIDVSAISNLI 1187
Qy 1083 QHWVPGARLVEELPHELVLVLPYTGADHSGSFATLFRBLDTRLAELRLTYGIGSDTSLEE 1142
Db 1188 RKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRLSDLGISYIGSETTLEE 1247
Qy 1143 FLKVVCEADTDMEDSCCGHLCTGIAGLDVTLRLKMPQETALE--NGEPAGSABETD 1200
Db 1248 FLKVAESGVDAETSDGTLFARRNRRAFG-DKOSCLRPFTEDDAADPNDSIDIPESRETD 1306
Qy 1201 QGSGPDVAG--RVQGWALTQOQLALLKRFLARRRRGLFAQIVLPALFVGLALVFSL 1258

RESULT 2

US-09-032-438C-118
; Sequence 118, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir

Db 1307 LLSGMDCKSGYQVKGWKLTOQQFVALLWKLLIARRSRKGFPAQIVLPAVFCIALVPSL 1366
Qy 1259 IVPPFGHYPALRISPTMYGAQVFFSADAPGDPGRARLLLEALLQEAG-----LEE 1308
Db 1367 IVPPFGKYPSLELQPMWYNEQYTFVNSNADPEDTGLLELLNALTKDPGFGTRCMEGNEFIPD 1426
Qy 1309 PPVQSHRSRFAPEVPAEVAKVLASGNWTPESSPACQSQPGARRLLPPCPAAGGPPP 1368
Db 1427 TPCQAGEEETTAPVQTIMDLFQNGNWTMQNPSPACQSSDKIKKMLPVCPPGAGGLPP 1486
Qy 1369 POAVTSGEVVQNLTRNLSDFLVKTYPRLVRGLKTKWNVVRYGGSFSLG-GRDGPLP 1427
Db 1487 PQKONTADTLQDLTGRTNISDYLVTYQIIAKSLKMKIWNREFRYGGSFSLGVSNTQALP 1546
Qy 1428 SGQELGRSVEELWALLSPFGGALDRVLKNLTAWAHSLODQSLKIWNFNKGWHSVAVF 1487
Db 1547 PSQEVNDIAIKQMKHKLKAKDSSADRLNSLGRFMTGLDTRNNVKNVFNKWHIAISSFL 1606
Qy 1488 NRASNAILRAHLPPGPARHAHSITTLNPLNLKEQLSEAAALMASSVDVLVSVICVFPMS 1547
Db 1607 NVINNAILRANLQGENPNSHYGITAFNPLNLTKQQLSEVALMTTSVDVLVSVICVFPMS 1666
Qy 1548 FVPASFTLVLIBERTVTRAKHLQMLMGLSPTLYWLGNFMDMNCVLPACIWWLIFLAFQO 1607
Db 1667 FVPASFVFLIQRVSKAKHLQFISGVKPVYIWSNFWMDMNCVVPATVILIFCFOO 1726
Qy 1608 RAYVAPANLPAALLLLLLGWSITPLMPYSPFPSPSTAYVVLVTCINLFIGINGSMATF 1667
Db 1727 KSYVSSNLPVALALLLLGWSITPLMPYSPFVKIPSTAYVVLVTSVNLFIGINGSVATF 1786
Qy 1668 VLELFSQDKLOEVSRLKQVFLFPHFCLGRGLIDMVRNOAMADAPERLDRQFQSPLRW 1727
Db 1787 VLELFTDKLNNLNDILKSVFLFPHFCLGRGLIDMVRNOAMADALERFGENRFPVPSLW 1846
Qy 1728 EVYGNKLAMVIOGPLFLLFTLLQHRSQLLPQVRSLPLLGEEDEADVARERERVVOGA 1787
Db 1847 DVVGNLFAVAVGVVFLITVLIOYRFFRPPVNAKLSPLNDEDEADVARERQRILDGG 1906
Qy 1788 TQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGCEGFLGNGAGKTTSTFRMVTGDTLAS 1847
Db 1907 QONDILEIKELTKIYRRKRKPAVDRIKCVGIPGCEGFLGNGAGKSSSTFKMLTGDITVT 1966
Qy 1848 RGAVALAGHSVAREPSNAHLSMGCYQSDAIFELLTGREHLELLARLGRVPEAQVQTAG 1907
Db 1967 RGAFLNKNLSILNIHEVHQNMGYCPQDAITELLTGREHVEFFALLRGPVEKEVGKVE 2026
Qy 1908 SGLARGLSYADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLWNS 1967
Db 2027 WATRKGLVKYGEKAGNYSGGNKRKLSTAMALIGPPVFLDEPTTGMDPKARRFLWNC 2086
Qy 1968 LLAVVREGSRVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTURVP 2027
Db 2087 ALSVWKEGRSVLTSMSMECEALCTRMAIMVNGRFRCLGSGVQHLKNRFGDGYTIVVRIA 2146
Qy 2028 AARS--QPAARFAVAAFPGSELPEAHGRLRFPOLPGRCALARVCELAHVGAHEVED 2085
Db 2147 GSNPDLKVPQDFGLAPPGSVLKKEKRNMLQYOL-PSSLSLARIFISILSQSKRLLHED 2205
Qy 2086 FSVSQTMLEEVLYFSGKQCKDE---DTEQKEAGVGDVPAPGLQHPKRVSOFLDDPSTA 2142
Db 2206 YSVSQITLQOVFNFAKQDQSDDDLKXLSLHKQTV-VDVAV-----LTSFLQDEKVK 2257
Qy 2143 ET 2144
Db 2258 ES 2259

APPLICANT: Sun, Hui
APPLICANT: Lupeki, James R.
APPLICANT: Nathans, Jeremy
APPLICANT: Anderson, Kent L.
APPLICANT: Leppert, Mark
APPLICANT: Dean, Michael
APPLICANT: Singh, Nanda
APPLICANT: Shroyer, No. 6713300h F.
APPLICANT: Smallwood, Philip M.
APPLICANT: Allikmets, Rando
APPLICANT: Lewis, Richard A.
APPLICANT: Li, Yixin
TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
TITLE OF INVENTION: ATP-Binding Cassette Transporter
FILE REFERENCE: BYLR-0065
CURRENT APPLICATION NUMBER: US/09/032,438C
PRIORITY FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: US 60/039,388
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.2
SEQ ID NO 118
LENGTH: 2261
TYPE: PR1
ORGANISM: Mouse
US-09-032-438C-118

Query Match 51.6%; Score 5747.5; DB 4; Length 2261;
Best Local Similarity 49.8%; Pred. No. 0;
Matches 1146; Conservative 370; Mismatches 588; Indels 195; Gaps 27;
QY 1 MAFWTQLMLLWKNFMYRRQVQLLVLLWFLFLVAVRHSHPPLEHHECHFPNKP 60
DB 1 MACKQRLMLLWKNLTFRRQTCQLLEWAWFLFLILSVLRLSYPPYQHECHFPNKA 60
QY 61 LPSACTVPMQLGICNVNNTCPQLTGPEEGRLSNFNDLSVRLADARTVLGASAH 120
DB 61 MPSAGTLFWQGIICVANNPCRYPTGEGAGVGNFNKISVSLFSDAQRLLLYSQDRT 120
QY 121 TLAGLGLKIATRAAR-----STAQOQTKQSPLEPM--- 153
DB 121 SIKDMHKVRLMLRQIKHPNSNLKQDFLVDNFTSGFLQHNLSLPRSTVDSLQKNGVLQ 180
QY 154 -----LDVAELL----- 160
DB 181 KVFQGYQLHLASLNCNGKLEBIIQLGDAEVSALCGLPRKLDAAERVLRYNMDILKPVV 240
QY 161 -----TSLRTESLGLAQOEPPLHSLLEAEDLAQELLALSLVLEL---ALIQPRG 212
DB 241 TKLNSTSHLPTQHLA-----EATVLLDSLGLAQELFSTKSWSDMRQEWMLTNVNS 293
QY 213 TSGPLEL---LSEALCSVRGSSSTVGPSNLNWEASDLMLVG-----QEPESALPDSSLS 265
DB 294 SSSSTQIYQAVSRIVCGHPEGGGLKIKSLNMYEDNNYKALFGGNNTEBDVDTFYDNSTTP 353
QY 266 ACSELIGALDSHPLSLRLWRLKPLILKLLPAPDTPFTKMLAQNVTSEELTLRLDVR 325
DB 354 YCNDLWKNLESPLSRITKWLKPLVLGVKILYTPATROVMAENVNTFOELAVFHOLE 413
QY 326 EWEMLGPRIFTFMDNDSNVAMLQRLQWQ--DEGRQPRPGRRDHMEALRSFL----- 377
DB 414 GWEELSFIQWTFMENSQEMDLVRLTLLDSRGNDQFEQKLDGLDWTADQIMAFKLNPED 473
QY 378 --DPGSGYSQODAHVGHVLTGRVTECLSDKLEAAPSEAAVLSRAQLLAEHFW 435
DB 474 VQSPNGSVYTWREAFNETNOAIQTISREFMECVNLNKLPIPTFVRLINKSMELDERKFW 533
QY 436 AGVFLGPDSSDPTHEPTDLG--PGHVRIRKIMDIDVVTIRTKIRDFWDPCGAADPL 493
DB 534 AGIVFTG-----ITPDSVELPHHVKKIRKIMDIDNVERTNKIKDGYWDPCGRADPF 583
QY 494 TDLRYVWGGFVYQLDLVERAAVRLSGANPRAGLYLQOMPYPYVDDVFLRLVLSRLPLF 553

DB 584 EDMRYVWGGFAYLQDVVEQAIIRVLTSSEKKTGVVYQQMPYCYVDDIIFLRVMSRMLPF 643
QY 554 LTLAWIYSVTLTKAVVREKETRLRDLTDWRMGLSRAVLWGLFSLCLGFLISAALLV 613
DB 644 MTLAWIYSVAIIKSIIVYKEARLKTMRIMGLDNGILWFSWFSVSLIPLLSAGILLV 703
QY 614 LKLGDIPLVSHPGVVFLFAAFVAVATVTSFLLSAFFSRANLAAACGGGAYLSLYPYVL 673
DB 704 LKLGNNLLPYSDPSVVFVFLSVFAMVLIQCFILISTLFRANLAAACGGIYFTLYPYVL 763
QY 674 CVAMRDLRIPAGRVAASSLSPVAFGFCESLALLEEQEGEGAHNVNTRPT-ADVFSLAQ 732
DB 764 CVAMQDYVGSIKIFASLLSPVAFGFCYFALPEEQIGVQWMDNLFSPVEEDGNLT 823
QY 733 VSGLLLLDAALYGLATWLEAVCPQYGIPEPWNPFRSRYWCGPRPKSPACPTPLD- 791
DB 824 AVSMMLFDFTLYGVNMTWYIEAVFPQYGIIPRWPYFCTKSYWFGBEIDEKSHPSQKGV 883
QY 792 PKVLVEEAPGLSPGVSVRSLEKRPFGSPOPALRGLSLDFYQGHITAFIAGHNGAGTTTL 851
DB 884 SEICHEEPTHLRLGVSIONLVKVRDGMKVAVDGLALNFYEGQITSLFCHNGAGKTTM 943
QY 852 SILSGLFPSPGGSAFILGHVRSMAAIRPHLGVCPQYNVLFDMLTVDLTVDEHVMFYGLKGL 911
DB 944 SILTGLFPPTSGTAVILGDIRSEMSSIRQNLGVCPQHNVLFDMLTVEEHIFYARLKG 1003
QY 912 SAAVGPPEQDRLLQDVL--VSKQSVQTRHLGGMORKLSVAIAFVGGSGOVILDEPTAGV 970
DB 1004 SEKHVKAEMEQLADVGLPSPKLSKTSQSGMORKLSVALAFVGGSKVWILDEPTAGV 1063
QY 971 DPASRRGIWELLKYREGRTILSTHLDLAEELGDRVAVAGGRLCCGSLFLRRHGL 1030
DB 1064 DYSERGIWELLKYRQRTILSTHMDLDEADILGRIATISHGKLCVCGSSSLFKNQGL 1123
QY 1031 SGYYTLTKARLPLTTN-----EKADTDMEGSVDTROBKNGSQSRSVGTPO 1078
DB 1124 TGYTLTKVKKDVSSSLSCRNSSSTVSLCKKEDSVSQSSSDAGLSGDHESDTLTIDVSAI 1183
QY 1079 LALVQWVPGARLVEELPHELVLVLPYTGANDGSPATLRFRELDTHLAEURLTGYISDTS 1138
DB 1184 SNLRKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRSLDSLGSIYSISET 1243
QY 1139 LEEFLKVVESCAADTDMEDG-----SCQHLCTGIAGLDVTLRLKMPQ 1183
DB 1244 LEEFLKVAESGVDAETSDGTLPARNRRAFGDKQSC-LHPFTEDDAVD-----PN 1294
QY 1184 ETALENGPAGSAPETDQSGSPDVG--RVQGWALTROQLQALLKREFLLARRSRGLFA 1241
DB 1295 DSDID----PESRETDLLSGMDGKGSYQLKGWKLTKQQFVALLWKRLLIARRSRKGFPA 1349
QY 1242 QIVLPALFVGLALVPSLIVPPFGHYPALRLSPTMYGAQVSFESEDAPODGPGRILLEALL 1301
DB 1350 QIVLPAPVFCIALVPSLIVPPFGYPSLELQPMWYNEQYTFVSNDAPEMDGTMQELNALT 1409
QY 1302 QEAG-----LEBPPVQHSRRPSAPVPAEVAKVLASGNWTPSPSPACQSQPG 1351
DB 1410 KDPGFGTRCMGNPIPDTPCLAGEEDWTISVPQISIVDLFQNGNWTMKNPSPACQSSDK 1469
QY 1352 ARRLPDCPAAAGGPPPPQAVTVGSGEVQNLTGRLNSDLVKTYPRLVQGLTKTKWYNE 1411
DB 1470 IKKMLPVCPGAGGLPPQPKQKTADILQNLTGRLNSDYLKTVQIIAKSLKNKIWYNE 1529
QY 1412 VRYGFSLG-GRDPGLPSGQBLGRSVEELWALLSPLPGALDRVLKNTLTAHSLDAQDS 1470
DB 1530 FRYGFSLSGVNSQALPPSHEVNDIAIKQMKLLKLTQDTSADRFLSSLSGRFMAGLDTKN 1589
QY 1471 LKIWFNKGWHSVAVFVNRASNAILRAHLPGPARHAHSITTLNHLPLNLTKEQLSEALM 1530
DB 1590 VKWFWNKGWHAISFLNINNALIRANLQGENPSQYGIITAFNHLPLNLTQQLSEVALM 1649
QY 1531 ASSVDVLVSIQVWPFAMSPASFTLVLFIEERTVRAKHLQMLMGLSPTLTYWLGFLNLMCN 1590

```
Db 1650 TTSDVULVSCVIFPAMSFVFPASVFLIQRVSKAKHLQFISGVKPIVYMLSNFVWDMCN 1709
Qy 1591 YLVPACTVVLIFLAFOQRAYVAPANLPALLLLLLLYGWSITPLMYPASFVFFSPSTAYV 1650
Db 1710 YVVPATVLIIFIGFQKQSVSSNLPVALLLLLLLYGWSITPLMYPASFVFKIPSTAYV 1769
Qy 1651 LTCINLFIGINGSWATVLELFSQKQVSRILKQVFLFPHPFCIGRGLIDWVRNOAMA 1710
Db 1770 LTSVNLFIGINGSVATVLELFTNNKLNINDILKSVFLFPHPFCIGRGLIDWVRNOAMA 1829
Qy 1711 DAPERLGDROFQSPRLREVGKNNLAVIOGFLPFLFTLLQHSQLLPQPRVRSPLLG 1770
Db 1930 DALERFGENFVSPSLWDLVGRNLFAMAVGVVFFLTVIQYRFFRPRPVKAGPLN 1889
Qy 1771 EEDDVARERERVQGGATQGDVLVRLNLTQYRGORMPAVDRLCLGIPPGCECGLLGNG 1830
Db 1890 DEDEDVRRERQRIIDGGQNDLIEIKELTKYRRKRPVADRICIGIPPGCECGLLGNG 1949
Qy 1931 AGKSTTRMTVGTDLASRGAVLAGHSVAREPSNAHLSMGYQPSDAIFELLTGREHLEL 1890
Db 1950 AGKSTTFKMLTGTDFVTRGDAFLKNLSILSNIHEVHQNMGYCQPFDAITELLTGREHVEF 2009
Qy 1991 LARLGVPEAQVACTAGSGLARLGLSMVADRPACTYSGGNKRKLATALALVGDPAVVFLD 1950
Db 2010 FALLRGVPEKEVGKFGWEARKLGLVKYGEKIASNYSGGNKRKLSTANALIGPVPVFLD 2069
Qy 1951 EPTTGMDPSARRFLWNSLLAVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQ 2010
Db 2070 EPTTGMDPKARRFLWNCALSIVKEGRSVLTSMSMECEALCTRMAIWNNGRFRCLGSPQ 2129
Qy 2011 HLKGRFAAGHTLTLRVPAARS--OPAAAFAVAEPFSGSELREAGGRURFOLPGGRCALA 2068
Db 2130 HLKNRFGDGTIVVRIAGSNPDLKPVQEFFGLAPFGSVLKEKHNMLQYQL--PSSLSSLA 2188
Qy 2069 RVFGLAVHGAHEGVDFSVSOTMLBEVLYFSKQCKDE--DTEQKEAGVGVDPAPG 2125
Db 2189 RIFSILSQSKRLHIEDIYSVSTTLOQVFNFAKQDDSHLKDLSLHKQTV--VDVAV- 2246
Qy 2126 LQHPKRVSQFLDDPSTAE 2144
Db 2247 -----LTSFLQDEKVKES 2259

RESULT 3
US-09-032-438C-3
; Sequence 3, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratneter, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2273
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-3
Query Match 44.3%; Score 4936; DB 4; Length 2273;
Best Local Similarity 44.3%; Pred No 0;
Matches 1029; Conservative 367; Mismatches 546; Indels 282; Gaps 39;
Qy 1 MAFVQTQMLLLWKNFMYRRRQPVQLLVLLWPLFLFVLVVRHSHPPLEHHECHFPNKP 60
Db 1 MGFVRQIQLLWKNWTLRKQKIRFVVELVWPLSLFLVLIWLNANPLYSHHECHFPNKA 60
Qy 61 LPSAGTVPMLOGLICNVNNTCFPOLTPGEBPGRLSNFNDLSVRLADARTVLGGASAH 120
Db 61 MPSAGMLPMLOGLIFCNVNNPCFQSPGSPGVSVNNSILARVYRDFOELLNAPESQ 120
Qy 121 TLAG-----LGLKLTATLRA----- 134
Db 121 HLGRIWTELHILSQFMDTLTHPERTAGRIRDIRDILKDEETLTLFLIKNIGLSDSVYL 180
Qy 135 -ARSTAQOPQTKOSPLEPPMLDVA----- 157
Db 181 LINSQVRPEQFAHGVDPDLAKDIACSEALLERFIIFSQRRAKTVRYALCSLSQGLQWI 240
Qy 158 -----ELTSLRTESLGLAQOEPHLSLEAAEDLAQELLALRLSVEL 203
Db 241 EDTLYANVDFKLFVRVLTLLDSRSQGINL-----RSGWGGILSDMSPRQIEFIHRPSMODL 296
Qy 204 ----RALLQR--PRGTSGLLELSEALCSV-RGPSSTVGPSPLNWYASDLMELVGQSPES 256
Db 297 LWTRPLMQNGGETTKLGMILSDLLCGYPEGGGSRV-LSFNWYEDNNYKAFGLIDSTR 355
Qy 257 ALP-----DSSLSPACSELIGALSHPLRLMLRLKPLILGLKLLFAPDPTFTTKLMAQVN 312
Db 356 KDFIYSYDRRTTSCFNCALIQSLNESNPLTKIAWRAAKPMLMGKILYTPDSPAAARLKNAN 415
Qy 313 RTEELTLRLDVRVEMGLGPRIFTFMNDSSNVAMLQRL-----QWQDEGRQ 361
Db 416 STFEELHVRKLVKAWEEVGPIWYFPDNTQNMIRDTLGNPTVKDFLNRQGEIGIT- 474
Qy 362 PRPGRDHMEALRSFLDPG-----SGYSQMODAHADVGLVGLVTECLSLDKLE 413
Db 475 -----AEALNFLYKGPRESQADDMANFDRDIFNITDRTLRLVNVQVLECLVLDKFE 526
Qy 414 AAPSEALVSRALQOLAEHRFWAGVFLGPEDSSDPTHEHTPDLGPGHVRKIRKMDIDVV 473
Db 527 SYNDETQLTORALSLEENFWAGVVF-----PDWYPTWSSLPHPVKIRKMDIDVV 578
Qy 474 TRTNKIRDREWDPGPADPLTLRLYVWGGFVYLQDLVERAAVRLSGANPRAGLYLOMP 533
Db 579 ETKNKIKORYWDSGPRADPVEDFRIYWGFPAYLQDMVEQGITRSQVQAEAPVGIYLOMP 638
Qy 534 YPCYVDDVFLRVLRSRLPLTLAWIYSVTLTKAVVREKETRLRDTMRAMGLSRVYLWL 593
Db 639 YPCFVDDSFMIILNRCFPFIFWVLAWIYSVMTVKSIVLEKELRLKETLKNQGVSNVAVIC 698
Qy 594 GWFSLCGLPFLLSAALLVLVLKGDILPYSHGPGVVFLLAAFAVATVTSQFLLSAFSPRA 653
Db 699 TWFLDSPSINMSIFLLTIFIMHGRILHYSDPFLFLFLAFASTATIMLCFLSTFFSKA 758
Qy 654 NLAACGLLAYFSLYLYPVLCVAVRDLRPLAGGVAASLLSPVAFGFCESLALLEEGEG 713
Db 759 SLAAACSGVIYFTLYPHILCLFAWQDRMTAELKAVSLLSPVAFGFTYLVFEEGLG 818
Qy 714 AQWHNVGTTRPT-ADVPFSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPNWFFPFRS 772
Db 819 LQNSIGNSPTEGDEFSFLLSQMOMLLDAAVYGLLAWYLDQVFPDGYGTPLPWFYLLQES 878
Qy 773 YWCGPRPKSPAPCPT-----PL-----DPK-----VLVEEAPGSLSPGVSVR 810
Db 879 YWLGG-----GCSTREERALEKTEPLETEEDPEHPEIGHDSFFEREHPGWVPGVCVK 932
Qy 811 SLEKRPFGSPQALRGSLDFYQGHITAFGLHNGAGKTTTSLISLGLFPFPPSGGSAFILGH 870
```

Db 933 NLVKIFPCGRPAVDRLNITFYENQITAFGLHNGAGKTTLSILTGLLPPTSGTVLVGR 992
Qy 871 DVRSSMAAIRHLGVCPOYNVFLDMLTVDHWHVFWGRKLGISAUVGPEQRLLODUGLV 930
Db 993 DIETSLDAVROSLGCPQHNLTFHHLTVAEHMLFYAQLKGSQEAQEMEAUMEDTGLH 1052
Qy 931 SKQSVQTRHLSGGMORKLSVAIAFVGGQVWILDEPTAGVDPASRRGIWEILLKYREGRT 990
Db 1053 HKRNEEAQDLGGQMOKLSVAIAFVGDAAKVVILDEPTSGVDPSRRSIWDLKLYRSRT 1112
Qy 991 LILSTHHLDABELGDRVAVAGRLCCGSPFLFRRHLGSGYYITLVKARLPLTNEKA 1050
Db 1113 IIMSTHHMDEADLLGDRITAIQAQGRLYCGSTPLFLKNCFGTGLYTLVR----- 1161
Qy 1051 DTDMEGSDVTQKKNKSGOG---SRVG-----TP-----QALLAVOHM 1085
Db 1162 -----KMNITQSRQKRGEGTSCSKGFSSTCPAHVDDLTPEQVLDGVDNEMDVLHH 1215
Qy 1086 VPGARLVEELPHELVLVLYTGAHDSFATLFRLELDTLRLAELRLTYGIGISOTSLEEFK 1145
Db 1216 VPEAKLVESICGELIFLJPNKFKHRAVASLFRLEETLADLGLSSFGISDTPLLEEFLK 1275
Qy 1146 VBECAADTMEGDSGC-----OHLCTG---IAGLDVTLRLKMPPOETALENGE 1191
Db 1276 VTE-----DSDSGPLFAGGAQKRENNVNRHPCLGPREKAG-----QTPQDSNVCSGA 1324
Qy 1192 PA-----GSAPETDQSGPDAGVRVQGWALTQOQLOALLKRPFLARRSRGLFAQIVLPA 1247
Db 1325 PAHPEGQPPPECEGPGOL---NTGQLVLQHVQALLVKRFQHTIRSHKDFLAQIVLPA 1381
Qy 1248 LFVGLALVFSLTVPPFGHYPALRSPMTYGAQVFFSEDAPDGRARLLLEALLQEAQ-- 1305
Db 1382 TFVFLALMLSIVIPFGYPALTLHPWYGGQYTFSSWDEFGSEQFTVLADVLNKKPGFG 1441
Qy 1306 -----LEBPPVOHSHRSAPVPAEVAKVILASGNWTPESPSPACQSQPGARRLLP 1357
Db 1442 NRCLKEGMLPEYPCGNST-PWKTPSVSNITQLOFQKQKWTQVNPSPSCRCSTREKLTMLP 1500
Qy 1358 DCPAAAGPPPOAVTGSVEVQNLTGENDFLVKTYPRLVROGLTKKWNVEVRYCGF 1417
Db 1501 ECPGAGLPPPTQRTQRTSLQDLTDNRISDFLVKTYPALIRSLSKSKFWNVEQRYGGI 1560
Qy 1418 SLGRDPLGLP-SGOELGRSVBELMALSLPLPGGALDR-VLKNLTAWAHSIDAQSLKIWF 1475
Db 1561 SIGGKLVPVITGEALVGLSDLGIMN-VSGGPITREASKEIPDFLKHLETDNIKWPF 1619
Qy 1476 NKGHWSVAFVNRASNAIRLAHLPFGPARHAHSITTLNHPNLTKESLSAALMASSVD 1535
Db 1620 NNKGHALVSLNVAHNAIRLASLPKDRSPBEYGITVISQPLNLTKESLSEITVLTTSVD 1679
Qy 1536 VLVSICUVFMSFVPASFTLVLIERTVTRAKHLQMGCLSPTLWGLNGLMDMGNLYLPA 1595
Db 1680 AVVAICVIFSFSFVPASFVLVLIQERNVNSKHLOFISGVSPPTVWNTFLMDIMNYSVA 1739
Qy 1596 CIVVLIFIAFOQRAYVAPANLPALILLLLYXWSITPLMYPASFPFSPSTAYVVLTCIN 1655
Db 1740 GLVVGIFIGFOKKAYTSENLPALVALLLLYGVAVIPMKYPASFLFVDPSTAYVALSCAN 1799
Qy 1656 LFIGINGMATVLELFSQDK-LOEVSRLKQVFLIPPHFCGLRGLDMVRQAMADAFE 1714
Db 1800 LFIGINSAITFILELPENRTLLRFNAVLAKLLIVFPHPFCGLRGLIDLALSQAQVDVYA 1859
Qy 1715 RLGRDQFQSPLRVEVGNLAKMVLQGLFLLFTLLLOHR---SOLLPOPRVRSLLPLGE 1771
Db 1860 RFGBEHSANPHWDLIGKFLFAMVVEGVVYFLLLVQRFHFLSQWIAEPTKE---PIV-D 1916
Qy 1772 EDEDVARERVRVQATOGDVLRLNLTKVYRGORMPAVDRLCLGIPGECFGLLVNGA 1831
Db 1917 EDDVAREORIIITGGNKTDILRLHELTKIYLTGSSPAVDRLCVGVRGECFGLLVNGA 1976
Qy 1832 GKTSTFRMVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELL 1891

Db 1977 GKTTFFKMLTGDTTVTSGDATVAGKSIILTNISEVHONMGYCPQFDAIDELLTGREHLYLY 2036
Qy 1892 ARLRGVPEAQVAQTAGSGLARLGLISWYADRPAQTVSGGNKRLATALALVGDPAVFLDE 1951
Db 2037 ARLRGVPAABETBKVANWSIKSLGLTVYADCLAGTVSGGNKRLSTAIALIGCPPLVLLDE 2096
Qy 1952 PTTGMDPSAREFLWNSLLAVVREGSVMLTSHSMBECEALCSRLAIMVNGRFRCLGSPQH 2011
Db 2097 PTTGMDPOARMLMNVIVISIREGPAVLTSHSMBECEALCTRLAIMVKGAFRCWGTIOH 2156
Qy 2012 LKGRPAAGHTUTLVRPAARS-----QPAAPFAAAEFPGSELREAHGRLRQLPFGGRC 2065
Db 2157 LKSPFGDGIYITMKIKSPKODLLPDLPNVEOFFQGNFPFSGVQRRHYNNMLQFV---SSS 2213
Qy 2066 ALARVFGELAVHGAHGVDEFSVQTMLEEVLYFSKDGKDED 2109
Db 2214 SLARIFQLLSHKSLLIEYSVTQTTLDQVFNFAKQOOTESH 2257

RESULT 4
US-09-032-438C-6
; Sequence 6, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupeki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 6
; LENGTH: 2235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-6

Query Match 43.6%; Score 4858; DB 4; Length 2235;
Best Local Similarity 44.0%; Pred. No. 0;
Matches 1019; Conservative 359; Mismatches 636; Indels 300; Gaps 40;

Qy 1 MAFVTQLMLLWKNFMYRRQPVQLLVLLWPLFLFVLVAVRSHHPLEHHECHFPNKP 60
Db 1 MGFVYRQIQLLWKNWTLRKQKIRFVVELWPLSLFLVLIWRNANPLYSHHECHFPNKA 60
Qy 61 LPSAGTVPWLQGLICNVNNTCFPOLTPGEEPRLSNFDSLVSRLLADARTVLGGASAH 120
Db 61 MPSAGMLPWLOGIFCNVNNPCFQSPFTGESPISVSNYNNLSLARVYRDFQELLNAPESQ 120
Qy 121 TLAG-----LGKLIATLRA----- 134
Db 121 HLGRIWTELHLSQFMDTLRTHPERIAGRIRDIRDKDEETLTKFLIKNIGLSVSVYL 180
Qy 135 -ARSTAPQPTKQSPLEPPMLDVA----- 157
Db 181 LINSQVRPEQPAHGVPDLAKDKIACSEALLERFIIFSQRGAKTVRYAKCSLQSTLOWI 240

APPLICANT: Landes, Gregory M.
 APPLICANT: Burn, Timothy C.
 APPLICANT: Connors, Timothy D.
 APPLICANT: Dackowski, William R.
 APPLICANT: Van Raay, Terence J.
 APPLICANT: Klinger, Katherine W.
 TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
 TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
 NUMBER OF SEQUENCES: 73
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENZYME CORPORATION
 STREET: One Mountain Road
 CITY: Framingham
 STATE: Massachusetts
 COUNTRY: United States of America
 ZIP: 01701
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/665,259
 FILING DATE: 17-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Dugan, Deborah A.
 REGISTRATION NUMBER: 37,315
 REFERENCE/DOCKET NUMBER: IG5-9.1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (508) 872-8400
 TELEFAX: (508) 872-5415
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1375 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-665-259-26

Query.Match 34.3%; Score 3818.5; DB 3; Length 1375;
Best Local Similarity 54.1%; Pred. No. 0;
Matches 754; Conservative 224; Mismatches 348; Indels 69; Gaps 14;

796	Qy	VEZAPGCLSPGVSVRSLEKRFPGSPOPALRGLSLDFVQGHITAPLGHNGAGKTTTTLSILS	855
		: : : : : : : : : : : : : : : : : : : : : : : : :	
2	Db	MEBEPHLRLGVSIQNLVLKVYRDMKVAVDGLALNFVEGQITSFLGHNGAGKTTTTMSILT	61
		: : : : : : : : : : : : : : : : : : : : : : : : :	
856	Qy	GLFPPSGGSFAILGHVDRSSMAAIRPHLGVCPQYNVLFMDLTVDEHWFVYGRLLKGLSAAV	915
		: : : : : : : : : : : : : : : : : : : : : : : : :	
62	Db	GLFPPTSTGYAIILGKOIRSEMSIQRNLGVCPQHNVLFMDLTVVEHIFWYARLKGUSEKH	121
		: : : : : : : : : : : : : : : : : : : : : : : : :	
916	Qy	VGPEODRLLODVL-VSKQSVQTRHLSGGNQMKLSVAIAFVGSGSQVVLDEPTAGVDPPAS	974
		: : : : : : : : : : : : : : : : : : : : : : : : :	
122	Db	VKAEMEQMALDVLGPPSKLKSQTSLSGMQMKLSVALAFVGSGKVVLDEPTAGVDPPS	181
		: : : : : : : : : : : : : : : : : : : : : : : : :	
975	Qy	RRGIWELLLKYRGRTLLIETHHLDENLLGDRAVVAGGRLLCCGCSPLFLRHLGSGYY	1034
		: : : : : : : : : : : : : : : : : : : : : : : : :	
182	Db	RRGIWELLLKYRGRTIILSTHMDENADILGDIRAIIISHGKLCCVGSLSFLKNQLGTTY	241
		: : : : : : : : : : : : : : : : : : : : : : : : :	
1035	Qy	LTLVKARLPLTTW-----EKADTDMGSDVDTQEKNGSGSGSVGTPQLLAV	1082
		: : : : : : : : : : : : : : : : : : : : : : : : :	
242	Db	LTLVKDVESSLSCRNSSSTVSLCKEDSVSSQSSDAGLGSDHESDTLTIDVSAISNLI	301
		: : : : : : : : : : : : : : : : : : : : : : : : :	
1083	Qy	QHWVPGARLVEELPHVLVLPYTGADHGSFATLFRELDTRLAELRLTYGIGSDTSLLEE	1142
		: : : : : : : : : : : : : : : : : : : : : : : : :	
302	Db	RKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDDRLSDLGTSSYGISETTLEE	361
		: : : : : : : : : : : : : : : : : : : : : : : : :	
1143	Qy	FLKVFECAADTMEDG-----SCGQHLCTGIGIAGLDVTLRLKMPPOETAL	1187
		: : : : : : : : : : : : : : : : : : : : : : : : :	
362	Db	FLKVAESGVDAETSDGTLPARNRRAFGDKQSC-LHPFTTEDDAVD-----PNDSD	412
		: : : : : : : : : : : : : : : : : : : : : : : : :	


```
Db 1361 --LTSFLODEKVKES 1373
RESULT 7
US-08-665-259-27
; Sequence 27, Application US/08665259
; Patent No. 6028173
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Backowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1457 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-665-259-27

Query Match 24.2%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 2.4e-231;
Matches 513; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 796 VEEAPPGLSPGVRSLSKRPFGSPQALRGLSLDFYQGHITAFIGHNGAGKTTLSLS 855
Db 1 MEEETHPLVVCVDKLTQVKNKDKLALNKLNLNLYENQVVSFLGHNGAGKTTMSILT 60

Qy 856 GLFPSPGSAFTLGHVDSMAAIPHLGVCPQYNVLEMDLTVDHWHFYGLKCLSAV 915
Db 61 GLFPPTSGSATYGHIDHTEMDKIRKNGMCPQHNVLFDRLTVEHLWFYSLKMSAQEE 120

Qy 916 VGPEQDRLLQDVLVSKQSVQTRHLSGGWQRKLSVAIAFVGSGSQVVIDDEPTAGVDPASR 975
Db 121 IRKETDKMIEDLSNKHSLVQTLSSGGMKRLSVAIAFVGSSRAILIDEPTAGVDPVAR 180

Qy 976 RGIWELLKLYREGRTLLSTHLLDBAELLGDRVAVAGRLCCCGSPFLFRHLGSGYYL 1035
Db 181 RAIWDLILKYPGRITLLSTHMDAELLGDRIAIISHGKLKCCGSPFLKGYXNDGYRL 240

Qy 1036 TLVKARLPLTTNEKADTDMGSDVTRQEKNGSQSRVGTPTQLLALVQHWVPGARLVEEL 1095
Db 241 TLVKQAPBPGTSQEPGLASSPGCPRL-----SSCSEPQVSOQIRKHAVASSLLVSDT 292

Qy 1096 PHELVLVLPYTCAGHDSFATLRFELDTRLAEILRTGYGISTLSREIFLKVVEE----- 1149
Db 293 STELSYIILPSEAVKGAERLFQOLEHSLDALHSSFGMLMDTTLLEVPFLKVEEEOQSLN 352

Qy 1150 CAADT-----DMEDGSCQHILCTGIAGLDVTLRLKMPPOETALENGEPAGSPETDQSG 1204
Db 353 SEADVKESSKDVLPAGELGTAVGGQAG-NLARCSELAQSQASQASSVSGSA-RGEBGTG 410

Qy 1205 -----PDVAG-----RVQGWALTQOQLQALLKRL 1230
Db 411 YSDGYGYRPLFDNLQDPNVLSQAEAMEALAQVQGSRKRGKMWLKMRFHGLVKRPH 470

Qy 1231 LARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPTRY-----GAQVSFF 1283
Db 471 CARRNSKALCSQILLPAFFVCVAMTVALSVEIGDLPPLVLSPSYQYHNYTQPRGNFIPA 530

Qy 1284 SED-----APGDPGRARLEALLQAG-----LEEP-----PVQHSSEH----- 1316
Db 531 NEERQEVRLRLSP-DASPOQLVSTFRLPSPGVGATCVLKSPANGSLGPMNLNLSGSSRLIA 589

Qy 1317 -RESAPEVPAEVAKVVLASGNWTPESPSP----- 1344
Db 590 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVXPDEDSLQAWNMSLPPTAGBETWTS 649

Qy 1345 -----CQCSPQAGARRLLIPDCPAAAGGPPPPQAVTSGSEVVQNLTRNLSDFL 1391
Db 650 APSLPLRVHEPVRCTCSAQGTGP---SCPSSVGGHPPQMRVV-TGDILTDTIGHNVSEYL 705

Qy 1392 VKTVPRLVRQGLTKKWNVEVRYGSGSLGGRRDPLPS--GQELGRSVBELWALLSPLPG 1449
Db 706 LFTSDRFLH-----RYGAITFGNVQKSIPTASFGARVPPMVRKI----- 744

Qy 1450 ALDRVLKNTAWAHSLDAQDSLKIFWNNKSHSVAFVNRASNAITLRAHLP-----GPARH 1506
Db 745 AVRRV-----AQVLYNNKGTHSMPTLYLNSLNNALTRANLPKSKGNPA-- 786

Qy 1507 AHSITTLNHPNLTKEOLESEALMASSVDVLVSVICVFPAMSFVPASFTLVITEERVTRAK 1566
Db 787 AXITVTNHPNKTSSASLDYLL-QGTDVVIATFIIIVAMSFVPASFVFLVAEKSTRAK 845

Qy 1567 HLQLMGGLSPTLYWLGNFMDMCMVLPACIVLVILFAQORAYVAPANLPAALLLLLY 1626
Db 846 HLQFVSGCNPIYWLANYVMDMLNLYLPATCCVILFVFDLPAYTPTNFPFVAVLSFLY 905

Qy 1627 GWSITPLMYPASFFFSVPSTAYVVLTCINLFIGNGMATFVLELFS--DOKLQEVSRILK 1685
Db 906 GWSITPIIMPASFWEFVFPSSAYVFLVINLFIGITATVATFLLQLFEHDKDKLVNSYLK 965

Qy 1686 QVFLIFPHFCLGRGLIDMVRNQAMADAPERLGD-RQFQSPLRWEVVGKNNLAWIQGDLF 1744
Db 966 SCFLIFPNYNLGHGLMEMAYNEYINEYKAKIQFDKMKSPFEWDIVTRGLVAMTVGFGV 1025

Qy 1745 LLFTLLQHRQLLPQPR---VRSPLLIGEDDEDVARERERVQVQATOGDVLVLRNLTKV 1801
Db 1026 FFLTIMCQY--NFLRQQLPQFVSTKPV--EDVDVASERQVRLRGDADNDVMVKIENLTKV 1081

Qy 1802 YRGO---RMPAVDRILCLGI-PPGECFGLGVNGAGKSTFRMTVGTDTLASGEAVLAGHS 1857
Db 1082 YKSRKIGRILAVDRCLGVCPVPGECFGLGVNGAGKSTFKMLTGDESTTGGEAPVNGHS 1141

Qy 1858 VAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSW 1917
Db 1142 VLKOLLVQQQSLGVCPCQFDVPDELTAHEHLQVTLRLRCIPWKDEAQVVKWALEKLELTK 1201

Qy 1918 YADRPAGTYSGNKRKLATALLVGDPAVFLDEPTTGMDSARRFLWNSLLAVVREGS 1977
Db 1202 YADRPAGTYSGNKRKLSTAILIGYFAFIDEPDTGMDPKARRFLWNLIDLTKGRS 1261

Qy 1978 VMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARS-OPAAA 2036
Db 1262 VMLTSHSMECEALCUTRLAIMVNGRLHCLGSIQHLKRGFGGYMITVTKSQSNQKDVVR 1321

Qy 2037 FVAEEFPGLSEUREAHGG-----RLRFQLPPGRCALARVFGELAVHGAHGVDFSVSQT 2091
```


Db 1322 FFNRFP-----EHAQKTPYQVQL-KSEHLSLAQVFSKMQVGVGLGIEDYSVSQT 1375
Qy 2092 MLEEVFLYFSKDGKQEDTERQKEAGVDPAP----GLQHPKRVSOFL-----DDPSTA 2142
Db 1376 TLDNVFNFAK--KQSDNVEQQAEPSSLPGLGSLLSLLRPRPAPTELRAVADEPDL 1432
Qy 2143 ET 2144
Db 1433 DT 1434
RESULT 8
US-08-762-500-27
; Sequence 27, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IG5-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1457 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-762-500-27
Query Match 24.2%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 2.4e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;
Qy 796 VESAPPGSLGCVSVRSLEKPPGSPQALRGLSLDFYQGHITAPLGHNGAGKTTLSILS 855
Db 1 MESEPTHLPLVVCVDKLTQVYKNDKLLNKLKSLNLYENQVSPFLGHNGAGKTTMSILT 60
Qy 856 GLPPPSGGSFILGHGDRSSMAAIRPHLGVCPQYNVLFDMLTVDHEWVFGRLKLSAAV 915

Db 61 GLFPPTSGSATIYGHDIRTEMDEIRKNLGMCPQHNVLFDRLTVEEHLMWFVSRLSKMAQEE 120
Qy 916 VGPEDQRLLDQVGLVSKSQSVQTRHLSGGMOKLSVAIAFVGGSQVWILDEPTAGVDPASR 975
Db 121 IRKETDKMIEDLESNKRHSLVQTLSCGMKRLSVAIAFVGGSEAILDEPTAGVDYAR 180
Qy 976 RGIWELLKLYREGRTLILSTHHLDEALLGDRVAVVAGGRLCCGSPFLRRHLGSGYYL 1035
Db 181 RAIWDLILKYKPGRTILLSTHMDADLLGRIATISHGKLCGCCSPFLKGYXDGURL 240
Qy 1036 TLAKARLPILTNEKADTMEGSDVTROEKNGSGSVGTPOALLALVOHVWPGARLYEEL 1095
Db 241 TLVKQPAEPGTSEPGGLASSPGCPRL-----SSCSEPQVQSQFIRKHWASSLLVSDT 292
Qy 1096 PHELVLVLPYTGARDGSAFATLFRLEDTRLAELRLTYGIGISDTLSLEEIFLKVVEE----- 1149
Db 293 STELSYILPSEAVKGAERLFOQLSHSLDALHLSFGLMDTTLVEFLVKVSEEDQSLN 352
Qy 1150 CAADT-----DMEDGSCGHLCTGIAGLDVTLRKMPPQETALENGEPAGSAPETDQSG 1204
Db 353 SEADVKEERKDVLPGAEGLTAVGQAG-NLARCSELAQSQASQSSVSGSA-RGERGTG 410
Qy 1205 -----PDVAG-----RVQGWALTROQLQALLKREL 1230
Db 411 YSDGYGYRPLFNDLPQDNVLSQEAEMEAALAVQGGSRKLEGWMLKORQPHGLLVKRFH 470
Qy 1231 LARRRRGLPAQIVLPALFVGLALVPSLIVPPPHGVPAALRLSPMTY-----GAQVSFF 1283
Db 471 CARRNSKALCSQILLPAFFVCVAMTVALSYPEIGDLPPLVLSQSYHNYTPQRCNFPYA 530
Qy 1284 SED-----APGDPGRARLLEALLQEAQ-----LEEP-----PVQHSSE 1316
Db 531 NEERQEVRLRLSP-DASPOQLVSTFRLPSGVGATCVLKSPPANGSLGPMNLSSGESLLA 589
Qy 1317 -RESAPEVPAEVAKVLASGNWTPESPSPA----- 1344
Db 590 ARFDSMCLESFTQGLPLSNFPPPPSPAPSDSPVXPDESLQANMSLPPTAGPTWTS 649
Qy 1345 -----CQCSQPGARLLPDCPAAAGPPPPQAVTGSVEVQNLTRNLSDFL 1391
Db 650 APSLPRLVHEPVRCTCSAQGTGF---SCPSSVGGHPQMRVV-TGDILTDTIGHNVSEYL 705
Qy 1392 VKTYPRLVROGLTKKWNNEVRYGGSLGGRDPLGS--QBELGRSVEELWALLSPPLGG 1449
Db 706 LFTSDRFLH-----RYGAITFGNVQKSIIPASFGARVPPMVRKI----- 744
Qy 1450 ALDRVLKNTLAWAHSILDAQSLKIWFNNKGWHSVAFVNRASNAIRLAHLPP---GPARH 1506
Db 745 AVRRV-----AQVLNNKGYHSMPTVYLSLNNAILRANLPKSKGNFA-- 786
Qy 1507 AHSITTLNHPNLTKBQLSEALMASSVDVLSICVVFAMSFPVSPASFTLVLIERVTRAK 1566
Db 787 AXITVTNHPMNKTSASLSLDYLL-QGTDVVIAIFIIVAMSFVPASFPVFLVAEKSTKAK 845
Qy 1567 HLQIMGSLSPTLVWGNFLWDMCNVLPACIVVLIFLAQORAVVAPANIPALLLLLLLY 1626
Db 846 HLQFVSGNCEVIYWLANYWMDMLNLPVATCCVILFVFDLPAYTSPTNPAVLSFLLY 905
Qy 1627 GWSITPLMYPASFFSPVSTAYVVLTCINLFIGINGSMATFVLELFS-DOKLGEVSRILK 1685
Db 906 GWSITPLMYPASFWFEVPSAYVFLVINLFIGITATVATFLQLQLEHDXDKLVNYSYLK 965
Qy 1686 QVFLIIPHFCLGRGLIDMVRNQAMADAFERLGD-RQFSPLRWEVVGKNIILAMVIOGPLF 1744
Db 966 SCFLIFPNYNLGHLMEMAYNEINEYAKIGQDKMKSPFEMDIVTRGLVAMTEGFGV 1025
Qy 1745 LLFTLLLOHRSQILLPQPR---VRSPLLGEDEEDVARERVVQGTQGVDLVRLNLTKV 1801
Db 1026 FFITIMCQY--NFLRQFQRLPFSVTKPV--EDDDVDVASERQVRLRGDADNDMVKIENLTKV 1081
Qy 1802 YRQO---RMPAVDRCLIGI--PPGECFGLLGVNGAGKTTSTFRMVTGDTLASRGEAVLAGHS 1857

Db 1082 YKSRKIGRILAVDRILCLGVCVPCGFCGGLLVNGAGKSTTFKMLTGDESTGCGEAFVNGHS 1141
Qy 1858 VAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARLCVPEAQVAQTAGSGLARIGLSW 1917
Db 1142 VLKDLQVQOQSGYCPQFDPVDELTAHRLQLYTRLCIPWKDEAQVVKWALEKLELTK 1201
Qy 1918 YADRPAGTYSGGNKRKLATALVGDPAVVFDEPTTGMDPGARRFLMNSLLAVVREGRS 1977
Db 1202 YADKPAGTYSGGNKRKLATAIAGVPAPIFLDEPTTGMDPKARRFLMNLIDLTKGRS 1261
Qy 1978 VMLTSHSMECEALSRALMNGRPRCLSGOHLKGRFAAGHTTLRLVPAARS-OPAAA 2036
Db 1262 VMLTSHSMECEALSRALMNGRPRCLSGOHLKGRFAAGHTTLRLVPAARS-OPAAA 1321
Qy 2037 FVAAPPGSELREAHGG-----RLRFOLPPGRCALARVFGELVHGAHGVDEFSVQOT 2091
Db 1322 FNRNFP-----EAAAGKTPYKQVQL-KSHISLAQVFSKWEQVGVGLGIEDFSVQOT 1375
Qy 2092 MLEEVLFPYKQKDEDETEQKAGVGVDPAP-----GLQHPKRVYSQFL-----DDPSTA 2142
Db 1376 TLDNVFVNFAPK---KQSDNVEQOEAPSSPLSPGLGLLSLLRPPAPTELRLVADEPEDL 1432
Qy 2143 ET 2144
Db 1433 DT 1434

RESULT 9
US-09-032-438C-119
; Sequence 119, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattnar, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupsaki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 1472
; TYPE: PRT
; ORGANISM: Mouse
US-09-032-438C-119

Query Match 24.2%; Score 2693.5; DB 4; Length 1472;
Best Local Similarity 40.8%; Pred. No. 3e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 796 VEEAPGSLSPGVNSLEKRPFGSPQALRGLSLDFYQGHITAFIGHNAGKTTLSLS 855
Db 16 MEEEPHTPLVVCVDKLTAKVNDKLAALKSLNLNENQVVSFLGHNGAGKTTTMSILT 75
Qy 856 GLFPSPGSAFILGHDFVSSMAAIPRHLGVCPQYNNVFLDMLTVDHVNWFYGLKGLSAAV 915
Db 76 GLFPPTSGSATIYGHDIRTEMDEIRKNLGMCPQHNVLPDRLTVEHLLWFLYRLKSMQEE 135

Qy 916 VGPEODRLLQDVGLVSKOSVQTRHLSGGMORKLSVAIAFVGGSOVVILDEPTAGVDPASR 975
Db 136 IRKETDKMIEDLELSNKHSLVQTLSSGGMKRLSVAIAFVGGSRAILLDEPTAGVDUPYAR 195
Qy 976 RGIWELLKRYREGRTLILSTHHLDBAELLDGRVAVVAGGRLCCGCSPLFLRHLHSGYYL 1035
Db 196 RAIMDLILKYKPGRTILLSTHMDADLLGDRIATISHGKLKCCGSPFLKGAJKDGYRL 255
Qy 1036 TLVKARLPLITNEKADTMEGSDVTRQEKNGSQSRVCTPQLLALVQHVHVPGARLVEEL 1095
Db 256 TLVKQAPALGSPQEGGLASSPGCPL-----SSCEPQVVSOFIRKHVASSLLVSDT 307
Qy 1096 PHELVLVLPYTGADHGFATLRELDTRLAELRTGYGISTDLSLSEIFLKVVEE----- 1149
Db 308 STELSIIIPSEAVKGAERLFOQLEHSLDALHLSFGMLDITLSEVFLKVSEEDQSLEN 367
Qy 1150 CAADT-----DMEDGSCQHILCTGIAGLDVTLRLKMPPOETALENGEPAGSPETDQSGG 1204
Db 368 SEADVKEKRDVLPAGEGLTAVGGQAG-NLARCSELAQSASQASASSVGSAR-GBEGTG 425
Qy 1205 -----PDVAG-----RVQGWALTQOLQALLKRL 1230
Db 426 YSDGYDYRPLPDNLQDPDNVLSQAEWALAAQVCGGRKLEGMWLKMRQFHLGLVKRRH 485
Qy 1231 LARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPMTMY-----GAQVSFF 1283
Db 486 CARRNSKALCSQILLPAFFVCVAMTVALSVPEIGDLPPLVLSFQYHNYTQPRGNFIPYA 545
Qy 1284 SED-----APGDPGRARLLEALLQBAQ-----LREP-----PVOHSSH----- 1316
Db 546 NEEROEYRLRLSP-DASPOQLVSTPRLPSPGVGATCVLKSPANGSLGMLNLSGSRLLA 604
Qy 1317 -RFSAPVPAEVAKVLASGNWTPSPSA----- 1344
Db 605 ARFEDSMCLSFSTOGLPLSNFVPPSPAPSDSPVKPDEDSLQANWMSLPPTAGBETWTS 664
Qy 1345 -----COCQOPGARRLLPDCPAAAGGPPPPQAVTGSGEVQVNLTRNLSDPL 1391
Db 665 APSLRLVHEPVRCTCSAQTGF-----SCPSSVGGHPQWVRVV-TGDILTDTIGHNVSEYL 720
Qy 1392 VKTYPRLVROGLTKKXWNVRYGFSIGGRDPGLPS--GOELGRSVBELMALLSPLPGG 1449
Db 721 LFTSDRFLH-----RYGAITFCNVQKSIIPASFGARVPPMWAKI----- 759
Qy 1450 ALDRVLKMLTAWHSLDQDSLKTFWNKNGHSMVAFVNRASNAITLRAHLPP---GPABH 1506
Db 760 AVRRV-----AQVDYNNKNGHSMPTYLNSLNNALIRANLPSKGNPA-- 801
Qy 1507 AHSITTLNHLNLTKEQLSEALMASSVDVLVSIQVFPAMSFVPASFTLVLEERVTRAK 1566
Db 802 AVKITVTNHPNKTSSASLDYLL-OQTDVIAIFIIVAMSFVPASFVVFLVAEKSTAK 860
Qy 1567 HLQMLGGLSPITYWIGNFLWDMCNVLPACTIVLIFLAFQORAYVAPANLPAALLLLLY 1626
Db 861 HLQFVSGCNPIVYMLANYVMDMLNLYVPATCCVILFVDFLPAYTSPTNFPFVLSLELY 920
Qy 1627 GWSITPLMYPASFPFESVPSTAVVLTICNLFITGMSMATFVLELFS-DOKLQEVSRILK 1685
Db 921 GWSITPIMYPASFPFESVPSTAVVLTICNLFITGMSMATFVLELFS-DOKLQEVSRILK 980
Qy 1686 QVLFIFPHFCILGRGLIDMVRNQAMADAFERLGD-RQFQSPRLRWEVVGKLLAMVQGPFL 1744
Db 981 SCFLIFPNYNLGHGLMAYNEYIYAKIQDPKMKSPFEDIVTRGLVAMTVGFGV 1040
Qy 1745 LLFTLLLOHRSQILPOPR---VRSPLLGEDEEDVARERVRVQOAGTODVLVLRNLTKV 1801
Db 1041 PFLTIMCOY--NFLRQORLPVSTKPV--EDVDVASERQVRVLRGDADNDMVKIENLTKV 1096
Qy 1802 YRGO---RMPAVDRILCLGI-PPGSCFGLLVNGAGKTTSTFMVVTGDTLASRGEAVLAGHS 1857
Db 1097 YKSRKIGRILAVDRILCLGVCVPGFCGLLVNGAGKTTSTFKMLTGDESTTGGEAFVNGHS 1156
Qy 1858 VAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLCVPEAQVAQTAGSGLARIGLSW 1917

Db 1157 VLKDLLQVQSLGCPDFVDELTAHQLYTRURCIPWKDEAVQVWALEKLELTK 1216
Qy 1918 YADRPACTYGGNKRKLTATLALVGDPAVVFLODEPTTGMDPSARRFLWNSLLAVVREGRS 1977
Db 1217 YADKPACTYGGNKRKLTALALIGYPAFIFLDEPTTGMDPKARRFLWNLILDLIKTGRS 1276
Qy 1978 VMLTSHMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARS-QPAAA 2036
Db 1277 VMLTSHMECEALCSRLAIWNGRLHCLGSIQHLKGRFGDGYMITVTRKSSQNVKDVR 1336
Qy 2037 FVAAEFGSSELREAHGG-----RLRFOLPGRCALARVFGELAVGAEHGVEDFSVSOT 2091
Db 1337 FFRNRP-----EHAQKTPYQVYQL-KSEHISLAQVFSKMEVQVGVILGIEDYSVSOT 1390
Qy 2092 MLEBVLFFSKDQKDBTBEQKAGVGDPA-----GLQPKRVQSFL-----DDPSTA 2142
Db 1391 TLDNVFNPAK---KQSDNVEQAEFSSLPGLGLLSLLRPRPAPTELRAALVADEPDL 1447
Qy 2143 ET 2144
Db 1448 DT 1449

RESULT 10

US-09-032-438C-120
; Sequence 120, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupeki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C
; PRIOR FILING DATE: 1998-02-27
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 120
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-120

Query Match 22.5%; Score 2507; DB 4; Length 1704;
Best Local Similarity 33.7%; Pred. No. 2e-214;
Matches 639; Conservative 314; Mismatches 630; Indels 314; Gaps 47;
Qy 307 LMAQVNTFEELTLRDVREWEMLGPRIFFTMNDSSNVLQRLQWQDEGRQPRGG 366
Db 9 LLLWKNVTQKRKVLTVLELFL---PLLF-----SGLIWLRLKIQSENVPNATIYPG- 59
Qy 367 RDHEALRSFLDPGSGGSWQ-----DAHADVGLVGLGRVTECLSLDKLEAAESEAALV 422
Db 60 -QSIQELPLFTFFPPPGDTWELAYIPSHSDAAKTVTETVRAUVMNM-RVRGFPSEKDFE 117
Qy 423 SRALQLAEHFWAGVFLGPESSDPTHTPDLPDGPCHVRIKIRMDIDVVTR-----475
Db 118 DYIRYDNCSVLAADVFEHFNHS---KEPLPLAVKYLRFSTRNMYMTQTGSFFLK 174

Qy 476 -----TNKIRDFWDPGPAADPLT-----DLRVWGGFFVYLODLVERAAVRLVSGAN--- 522
Db 175 ETEGWHITSLFPLFPNPGP-RETPSPDGEGPIREGFLAVQHAVDRAINMEYHADAATRQ 233
Qy 523 --PRAGLYLQOMYPYPCVDDVFLRVLSRSLPFLTLAWIYSVTLTKAVVREKETRLRDT 580
Db 234 LFORLVTYTKRFPYPPFIEDPFLVAIQYQLPLLLLSFTYTTALTIAARVVOEKERRLKEY 293
Qy 581 MRAMGLSRVAVLWGLFSLCLGPFLLSAALVALVLKLG-----DILPYSHPGVVFLFLAAF 635
Db 294 MRMMGLSSMLHWSAWFLPLFLLIAASFMTLLFCVKVKNVAVLSRSDPSLVLAFLLCF 353
Qy 636 AVATVTSQSLSPAFPRANLAACGLAYPSLPLVPLVAVWRDRLPAGGRVAASLLSPV 695
Db 354 AISTISFSEFVSTFFSKANMAAAGGGLYFPTIYFFVAPRYNMWMTLSQKLSCLLSNV 413
Qy 696 AFGFGESLALLEEQEGQWNVGTRPTA-DVFSLAQVSGLLLLDAALYGLATWYLEAV 754
Db 414 AMANGAQLGKFEAKGMGIQWRDLSLPVNVDDDFCFQVGLGMLLLDSVLGLVTWMEAV 473
Qy 755 CPQGYGIPFPWNPFFRRSYWCGPRPKSPAPCPTPLDPKVLV-----EEAPPLSGPSV 809
Db 474 FPGQFGVPQPYFFIMPSYWG-KPRAVAGKEEDSDPEKALRNEYFEAPEDLVAGIKI 532
Qy 810 RSLEK--RPPGSPQALRGLSLDFYQGHITAFIHLNGAGKTTTILSLGILFPSPGGSAPFI 867
Db 533 KHLKSVFRVGNKDRAAARDNLNLYEGQITVLGLHNGAGKTTTILSMLTGLFPPTSGRAYI 592
Qy 868 LGHDVRSSMAAIRPHLGVCPOYNVFLDMLTVDSHVMFYGRKLGLSAAVWVPEQDRLIQDV 927
Db 593 SGFEISQDMVQIRKSLGCLCPQHDILFDNLTVAEHLFYAQLKGLSRQKCEEVKQMLHII 652
Qy 928 GLVSKOSQVTRHLSGGMQRKLSVAIAFVGSQVVLIDPTAGVDPASRRRGIWELLKYRE 987
Db 653 GLEDKWNRSRFLSGGMRRKLSIGIALIAGSKVILIDPTSGMDAISRAIWDLLQKRS 712
Qy 988 GRTLILSTHLDDEAILGDRVAVVAGGRLLCCSPFLRRHLGSGYVLTIVKARLPITTN 1047
Db 713 DRTIVLTTHFMDADLLGDRIAIMAKGELQCGSSSLFKOKYGAGYHMTLVKE----- 765
Qy 1048 EKADTMEGSDVTRQEKNGSQSRVGTPOALLAVQHVPFCARLVEBELPHELVLVLYPTG 1107
Db 766 -----PHCNPEDISQLVHHHVENATLESSAGAELSFILPRES 802
Qy 1108 AHGDSFATLPRDLRLAEURLTGYGISDTSLEBEIFLKV----- 1146
Db 803 TH--RPEGLFAKLEKKQKELGIAFGASITTBEEVFLRVGKLVDSMDIOAIQLPALQYQ 860
Qy 1147 VEECAADTMEDSCGCHLCTGIAGLDVTLRKMPQETALENGEPAGSAPETDQSGSPD 1206
Db 861 HERRASDWAVDNLG-----AMDPSDGI GALIEER-----T 893
Qy 1207 AVGRVGNWALTROOLOALLKRLFLARRSRRLFAQIVLPALFVGLAVFLVPPFGHY 1266
Db 894 AVKLTGLALHCOQFWAMFLKKAAYSREWKAQVVLPTCVTLALLAINYSSELFD 953
Qy 1267 PALRLSPTMYGAQVSPFSDAPGDPGRARLLEALLQAGLEEPVQVSHSRFSAPEVPAE 1326
Db 954 PMLRLTLGEXGRVVPFSV-----PGTSQLGQQL-----SEH----- 985
Qy 1327 VAKVLASGNWTPSPSPACQSQPGARRLLPDCEAAGGPPPPQAVTGVSGEVVQNLTRN 1386
Db 986 -----LKDALQAEQ-----QEPREVUG-----D 1003
Qy 1387 LSDPLVKTYPRLVROGLTKKVVNRYGFSFGGRDPLGFSQGLGRSVEELWALLSPL 1446
Db 1004 LEFLI-----FRASVEGGGFN----- 1020
Qy 1447 PGGALDRVLKVLTAHSHLDAQSLKTIWNNKGHSMVAFVFNASNAIILRAHLPPGPAPH 1506
Db 1021 -----ERCL--VAASFRDVGERTVVALFNQAYHSPATALVVDNLLFK--LLCGP--H 1069

Db 654 LEDKNSRSFLSGMRKLSIGIALTAGSKVLLIDPTSGMDAISRRAIWDLQKQSD 713
Qy 989 RTILSTHHIDEALGDRVAVVAGRLCCGSPFLRRHLGSGYLLTVKARLPITTNE 1048
Db 714 RTIVLTHFDEADLLGDRITAMAKGBLQCCGSLFLKQYAGYHMTLKE----- 765
Qy 1049 KADTMEGSDVTQEKNGSGSRVGTPTQLLAVQHWVPCARLVEELPHELVLVPTGA 1108
Db 766 -----PHCNPEDISQLVHHVNPATLESSAGAEISFILPREST 803
Qy 1109 HDGSFATLPRDLRLAELRTGYGSDTSLBIFLKV-----V 1147
Db 804 H--RFEGLFAKLEKKQELGASFGASITTWEEVFLVGLVDSSMDIQAIQALPALQYOH 861
Qy 1148 EECADTDMEDGSCGQHLCTGIAGLAVTLRLKMPQETALENGEPAGSAPETDQSGPDA 1207
Db 862 ERRASDWAVDNLGC-----AMDPDGDIGALIEBER-----TA 894
Qy 1208 VGRVQGWALTRQOLQALLAKRFLARRSRGLFAQIVLPALFVGLALVFSLIVPPGHYP 1267
Db 895 VKLNTGLALHCQDFWAMFLKKAAYSREWQAAQVLVPLTCTVTLALLAINYSSELPDDP 954
Qy 1268 ALRLSPMYCAQYFFEDAPGPRARILDEALLQEALEBPVQVSHSRFSAPVPAEV 1327
Db 955 MLRLTLCGEYRTVVPSV-----PGTSOLGQL-----SEH----- 985
Qy 1328 AKVLASGNWTPESPSPACQSQCARLLPDCPAAAGPPPPQAVTGSGEVQNLGRNL 1387
Db 986 -----LKDALQAEQ--QEPREVLG-----DL 1004
Qy 1388 SDFLVKTYPLRVQGLKTKWNEVRYGFSGLGRDPLGSPGQELGSRVEELWALLSPLP 1447
Db 1005 EEFLLI-----FRASVEGGGFN----- 1020
Qy 1448 GGALDRVLKNTLAWHSLDAQDSLIKIFNKGWHSVAVFNASNATLRAHLPPGPARHA 1507
Db 1021 -----ERCL--VAASFRDVGERTVFNALFNQVHSPATALAVDNLFLK--LLCGP--HA 1070
Qy 1508 HSITTLNHP-----LNLTKBQLGEAALMASSVDVLVSVICVVFVAFSPASFTVLVLEEV 1562
Db 1071 -SIWVSFPQPSALQAQKQFNEG--RKGFQIALNL--LFAMAFLASTFSLAVSERA 1124
Qy 1563 TRAKHLQMGSLSTLYWLNGLWDMCNLYVPACIVVLVILAFQOORAYVAPANLPALLL 1622
Db 1125 VQAKHVQFVSGVHVASFWSALLWDLISFLIPSLLLVVPKAFDRAFTRDGHMADTLLL 1184
Qy 1623 LLLYGWITPLMPYSPFFSPVSTAYVVLTCINLFIGINGSMATFVLEL---FSDKLOE 1679
Db 1185 LLLYGWAILPLMYLWNEFFLGAATYTLTIFNLISGI-----ATFLVMTIMRIPAVKLE 1240
Qy 1680 VSRILKQVFLIFPHFCLGRGLDMVRN-----QAMADAFERLGRDQFQSP-LRWEV 1729
Db 1241 LSKTLDHVLVLPNHLGMAVSSPYENYETRYCTSEVAAYHCKYNIQYQENFYAWSA 1300
Qy 1730 --VGNLLAMVIOG-----PLFLFTLLQHRSSOLLQPPVR-----SLPLGGEDE 1774
Db 1301 PGVGRFVMAASGAYLILLFIETNLQRLGILCALRRRTLTLYTRMPVL-PEDQ 1359
Qy 1775 DVARERVRVQATQG--DVLVRNLTKVYRGQMP--AVDRLCIGIPPGCEFGLLGVN 1829
Db 1360 DVADERTRILAPSDSLHPLIILIKELSKYVE--QRPVLLAVDRLSLAVQKGECEGLLGFN 1418
Qy 1830 GAGKTSTFRMVGTPLASGEAVLAGHSVAREPSAHLSMGYCQSDAIFELTIGREHLE 1889
Db 1419 GAGKTTTFKMLTGESLTSGDFAFGCHRSISDVGVQRIGYCPQFDALLDHMTGREMLV 1478
Qy 1890 LLARLRGVPQAQVATAGSLARGLSWYADRPAQTYSGGNKRLATLALVGDPAVVEL 1949
Db 1479 MYARLRGIPERHIGACVENTLRLGLLEPHANKLVRTYSGGNKRLSTGIALIGEPVIFL 1538
Qy 1950 DEPTGMDPSARRFLWNSLAVVREGSRVMTLSHMECEALCSRLAIMVNGRPRCLGSP 2009
Db 1539 DEPTGMDPVARRLLNWTVAPARESGKAILIITSHMECEALCTRLAIMVQGGFKCLGSP 1598

Qy 2010 QHLKGRFAAGHTLTLLVPAARSOPA-----AAFVAAREFGSELREAHGGRRLRFLFOLPPGGRC 2065
Db 1599 QHLKSGFGYSRLAKVSEGQOQALEEPKAFVDLTFPGSVLEDEHQGMVHYHL-PRDL 1657
Qy 2066 ALARVGEVLAVGAEGHVEDFVSQTMLEVEFLYFS 2101
Db 1658 SWAKVGEILKAKKEKYGVDYDYSVQSLSLEQVFLSFA 1693

RESULT 12

US-08-665-259-25
; Sequence 25, Application US/08665259
; Patent No. 6028173
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.1
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-259-25

Query Match 22.4%; Score 2501.5; DB 3; Length 1684;
Best Local Similarity 33.8%; Pred. No. 6.1e-214; Indels 309; Gaps 46;
Matches 635; Conservative 304; Mismatches 628;

Qy 327 VWEMLGPRITFMNDSSNVAMLQRLQWQDEGRQRPGRGRDHEALRSFLDPGSGYSW 386
Db 6 VLELFPLFLF-----SGILLWLRKIQSENVPNTIYFG--QSIQELPLFFTFPPPGDTW 58
Qy 387 Q-----DAHADVGHVLTGLRVTECLSLDKLEAAPSAALVSRALQLLAEHRFWAGVFLG 442
Db 59 ELAYIPSHSDAAKAVTETVRALVINM-RVRGPFSEKDFEDYIRYDNCSSVLAAVVEH 117
Qy 443 PEDSSDPTHEPTDLPFGHVRIKIRMDIDVVVTNKI---RDRFWD-----PGPAD 491
Db 118 PFNHS---KEPLFLAVKYHLRFYTRRNWNTQTGSFFLKETEGWHTTSLFPLFPNDGPR 174
Qy 492 PLT-----DLRYVWGVGFVYLQDLIVERAAVRLSCAN-----PRAGLYLQOMPYPYCVYDDV 541

Db 175 ELTSPDGEPGVIRESFLAVQHAQVADRAIMEYHADAAATRLQFORLVTVTIKRPPYPPFIADP 234
Qy 542 FLRVLSRSLPLFLTLAWIYSVTLTAKAVVREKETRLDRTMRAMGLSRVAVLWGLFSLCLG 601
Db 235 FLVAIOYQLPILLISPYTALTARAVVQEKERLKYMMGLSMLHWSANFLFLFL 294
Qy 602 PFLSAALLVLVLKLG-----DILPYSHGVVFLFLAFAVATVTSQFLLSAPFSRANLA 656
Db 295 FLIIAASFWTLFCVKVKNVAVLSRDPSLVLAFLCFAISTISFSPWSTFFSKANNA 354
Qy 657 AACGGLAYFSIYLPVYLVAHREDRLPAGGRVAAASLLSPVAGFGCESLALIEEQEGAOW 716
Db 355 AAFGGLFFTYPIYFFVAPRYNMWTLQKLCCLSNVAMGAAQLCKFEAKGMGIOW 414
Qy 717 HNVGRPTA-DVFSLAQYSGLLLDAAALYGLATWYEAQVCGQYGIPEPMNPPFPFRSYWC 775
Db 415 RDLSPVNVDDPCFGQVLGMLLDSVLYGLVTWYEAQVCGQYGIPEPMNPPFPFRSYWC 474
Qy 776 GPRPKSPAPCPTPLDPKVLV-----EBAPPGLSPGVSVRSLEK--RFPSPQPALRGLS 828
Db 475 G-KPRAVAGKEEDSDPEKALRNEYFEAPEDLVAGIKIKHLSKVFRVGNKDRAAVRDLN 533
Qy 829 LDFYQCHTAFGLHNGAKTTTSLSGLFPSPGGSFAFILGHVDRSSMAAATRPHLGVCPQ 888
Db 534 LNLVEGQITVLLGHNGAKTTTSLMGLTGLFPPTSGRAYISGYEISQDMVQIRKSLGLCPQ 593
Qy 889 YNVLFDMLTVDHVMFYGRKLKLSAAVVGPEODRLQDVLGVSKQSVOTRHLSGQMOKL 948
Db 594 HDILFDNLTVABHLFYAQLKGLSKQKPEVKQMLHIIIGLEDKNSRSRFLSGMRRKL 653
Qy 949 SVATIAVGGSQVILDEPTAGVDPASRRGIWELLKRYEGRTLILSTHHLDEAELLGDV 1008
Db 654 SIGIALAGSKVLLDEPTSGMDAISRAIWDLLQKQKSDRTIVLTHFMEADLGDRI 713
Qy 1009 AVVAGRLCCCGSPFLRRHLSGYYLTLVKARLPLTNEKADTDMGSDVTRQEKNGS 1068
Db 714 AIMAGELQCCGSSLFLKQYKAGYHMTLVKE-----745
Qy 1069 QGSRVGTPLQALVQHVHVGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTLRLAELR 1128
Db 746 --PHCNPEIDISQLVHHVHPNATLESSAGAEISFILPRESTH--RPEGLFAKLEKKQKELG 801
Qy 1129 LTGYGSDTSLIEEFLKY-----VEECAADTDMEDGSCGQHLCCT 1167
Db 802 IASFGASITTEVEFLRVKLVDSMDIQAQLPALQYQHERRASDWAVDNLCG-----856
Qy 1168 GIAGLDVTLRLKMPDETALENGEPAQAPETDQSGDPDVGVRVQGWALTQQOQALLLK 1227
Db 857 -----AMDPDSDIGALIBER-----TAVKLTNTGLALHCQCFWAMFLK 894
Qy 1228 RPLARRSRRLGFAQIVLPALFVGLALVSLIVPPFGHVPALRLSPMYGAQVFFSEDA 1287
Db 895 KAAVSWREKWKVAAQVLPVLCVTLLALAINVSELFDPPMLRLTLGEGYRTVVPFSV--952
Qy 1288 PGDGRARLLBALQEALEAGLEPPVQHSRHSFAPEVPAEVKVASGNWTPESPSPACQC 1347
Db 953 ---PCTSQLGQQL-----SEH-----965
Qy 1348 SQPGARRLLPCPAAGPPPPQAVTSGSEVVQNLITGRNLSDFLVKTYPRVLROGLKTKK 1407
Db 966 -----LKDALQAG--QEPREVILG-----DLBEFLI-----989
Qy 1408 WNVRYGFSIGGRDGLPGSQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSIDA 1467
Db 990 FRASVEGGGFN-----ERCL--VAASFRDVE 1014
Qy 1468 QDSLKIFNNKGMHSMVAVNRASNAILRAHLPPGPARHAHSITTLNHP-----LNLTKY 1522
Db 1015 RTVVNALFNNQAYHSPATALAVVDNLLFK--LLCGP--HA-SIVSVNFPQPRSAQAKD 1069
Qy 1523 QLSEANLMASSVDVLVSTCVVPFAMSVFVSPFTVLVIEBRVTRAKHLQMLGGLSPTLYWIG 1582
Db 1070 QFNEG---RKGFDIALNL--LFAMAFLASTFSLAVSRVQAKHVQFVSGVHVASFWLS 1124

Qy 1583 NFLWDMCNLYPACIVVLIFLAFOORAYVAVANLPAIILLLLLLYGWSITPLMYPASFPFS 1642
Db 1125 ALLWDLISFLIPSLLLLVVFKAFDVRAFTRDGHMADTLLLLLLLYGWAIIPLMYLWNPFL 1184
Qy 1643 VPSTAYVVLTCINIFIGINGSMATFVLEL-----FSDQKLQEVSRILKQVFLIPPHFCLG 1699
Db 1185 GAATAYTRLATIFNLISGI-----ATFLMVTIMRIPAVKLEELSCTLHDHFLVLPNHCIGMA 1240
Qy 1700 LIDVNRN-----QAMADAFERLGDROFQSP-L-RWEV--VGKNLLAMVIOG-----P 1742
Db 1241 VSSFYENYETRYCTSEVAHAHYCKYNIQYQENFYAWSAPGVRFRVMAASGCAYLIL 1300
Qy 1743 LFLFTLLQLQHSOLLQPPVR-----SLPLLGEDEVDVARERERVQATQG---D 1791
Db 1301 LFLIETNLLQRLGILCALRRRTLTLYTEMPVL--PEDQDVADERTRILAPSPDSLHT 1359
Qy 1792 VLVRNLTQVYRGQMP--AVDRCLGIPPGECFGLLVNGAGKTSTFRWVTGDTLASRG 1849
Db 1360 PLIIKELSKVYE--QRVPLLAVDRLSLAVQKGEFCFGLLVNGAGKTSTFRWVTGDTLASRG 1418
Qy 1850 EAVLAGHSVAREPSAAHLSMGCYCPQSDAIPELLTGREHLELLARLGVPEAQVACTAGSG 1909
Db 1419 DAFVGGHRISSDVGVKVRQIRGICPQFDALLDHTMGTREMLVYARLGRIPERHIGACVENT 1478
Qy 1910 LARLGLSWYADRPAGTYSGGNKRKLATALALVGPVAVVFLDEPTTGMDDPSARRFLWNSLL 1969
Db 1479 LRGLLPHANKLVRTYSGGNKRKLSTGIALIGEPANIFLDEPSTGMDPVARRLLWDIVA 1538
Qy 1970 AVVREGSRVMLTSHMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPA 2029
Db 1539 RARESGKAIITSHMEECEALCTRLAIMVQGFCLGSPQHLKSKFGSGYSLRAKVQSE 1598
Qy 2030 RSQPA---AAFYAAEPGSELREAHGRLRFPQLPGRCALARVFGELAVHGAHGYED 2085
Db 1599 GQBEALBEFAFADLTTPFGSVLEDEHQGMVHYHL--PGDLSWAKVFGILEKAKYGYDD 1657
Qy 2086 FVSQTMLEEVLYFS 2101
Db 1658 YSVSQISLEQVFLSFA 1673

RESULT 13

US-08-762-500-25
; Sequence 25, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-5415
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-762-500-25

Query Match 22.4%; Score 2501.5; DB 3; Length 1684;
Best Local Similarity 33.8%; Pred. No. 6.1e-214;
Matches 635; Conservative 304; Mismatches 628; Indels 309; Gaps 46;

QY 327 VWEMLGPRIETFWNDSNVAMIQELQWQDEGRQRPQGGRDHWEALRSFLDPGSGYSW 386
DB 6 VLSEFLPLF-----SGILWLRLKIQSENVPNATIYPG--QSIQELPLFFTFPPPGDWTW 58

QY 387 Q-----DAHADVGHLVGLRVTECLSLDKLEAAPSEAAALVSRALQLIAEHRFAGVVFGLG 442
DB 59 ELAYIPSHSDAAKAVTETVRALVINM-RVGRFPSEKDFEDYIRYDNCSSVLAAVFEH 117

QY 443 PEDSDTEHTPDLPGLGHVRIKIRMDIDVTRTKI-----RDRFD-----PGPAD 491
DB 118 PFNHS---KEPLPLAVKYHLRFSTRNNYMTOTGSPFLKETEGWHTTSLFPLFPNPGPR 174

QY 492 PLT-----DLRYVGGFVYLQDLVERAAVRVLSGAN-----PRAGLYLQOMPICYDDV 541
DB 175 ELTSPDGEFGYREGFLAVQHAVDRAIMEYHADAATRQFLQRTVTIKRPPYPPFTADP 234

QY 542 FLRLVLSRLPLFLTLAMIYSVTLVKAVVREKETRLDTRMGLSRVLWGLWFLSCLG 601
DB 235 FLVAIQOLPLLLLSFTYALTARAVQEKERLKEYRMGLSLWLSAWFLFLPFL 294

QY 602 PFLLSAALLVLVILKLG-----DILPSHGPGVFLFLAFAVATVQSFLLSAFSSRANLA 656
DB 295 FLIIAASFMTLLFCVKVKPNVAVLSRDPSSLVLAFLLCFAISTISFSFMVSTFFSKANMA 354

QY 657 AACGLAYFSLYPLVLCVAVRORLPAGGRVAASLLSPVAFGCESLALLEOGEAOW 716
DB 355 AAFGGFLYFTYIPYFFVAPRYNMTLUSQKLCCLSNVAMAGAQDIGFEAKGMGIQW 414

QY 717 HNYGTRPTA-DVFSLAQVGLLLDAAALYGLATWYLEAVCPGQVGIPEPWNFFPRSYWC 775
DB 415 RDLSPVNVDDDCFCGQVGLMLLDLSVLYGLVWYMEAVFPGQGVQVPPHYFIMPSYWC 474

QY 776 GPRPPKSPAPCPTPLDPKVLV-----EAPPGLSPGVSVRSLEK--RFPSPQAPALRGLS 828
DB 475 G-KPRAVAGKEEDSDPEKALRNEYFAEPEDVLVAGIKIKHLSKVFVRGNKDRAAVRDLN 533

QY 829 LDFEQGHITAFLGHNGAGKTTLSLSGLPPSGSAFILLGHVDRSSMAAIRPHLVCPO 888
DB 534 LNLVEGQITVLLGHNGAGKTTLSMLTGLPPTPSGRAIYSIGEYISQDMVQIRKSLGLCPQ 593

QY 889 YNVLFDMLTVDHFWVFYGRILKGLSAAVVGPEQDRLLQDGLVSKVSQVTRHLSGGMOKL 948
DB 594 HDILFDNLTVAEHLVFFAQLKGLSRQKCPPEVKQMLHIIGLEDKWNRSRFLSGMERKL 653

QY 949 SVIAIAFVGGQVILDEPTAGVDPASRRGIWELLKYREGRTILSTHHLDEAELLDGRV 1008
DB 654 SIGIATAGSKVLILDEPTSGMDAISRAIWDLIQROKSDRTIVLTTHFMDDEADLLGDRI 713

QY 1009 AVVAGGRLLCCGSPFLFRRHLGSGYYILTVKARLPLTTNEKAUDDMEGSDVTRQEKNGS 1068

DB 714 AINAKGELQCCSSSLFKOKYAGYHMTLVKE----- 745
QY 1069 QGSRVGTQOLLALVQHWVFCARLVEBELPHELVLVLPYTGADHGSFATLPRELTRLAELR 1128
DB 746 --PHCNPEDISQLVHHVFNATLESSAGAELSFILPRESTH--RFEGLPFAKLEKKQKELG 801
QY 1129 LTGYGISTSLBEIFLKV-----VEECAADTDMEDGSCGQHLC 1167
DB 802 IASFGASITTEERVFRLRVGKLVDSMSMDIQAIQLPALQYQHERRASDWDVSNLCG----- 856
QY 1168 GIAGLDVTLRLKMPPOETALENGEPAGSAPETDQSGDPDVGVRVQGMALTRQOLQALLK 1227
DB 857 -----AMDPSDGIGALIEER-----TAVKUNTGLALHCCQCFWAMFLK 894
QY 1228 RFLLAARRRRGLPAQIVLPALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVSFFSEDA 1287
DB 895 KAAYSRWKMWAAQVLPVLTCTVTLALLAINYSSELFPDDPMLRLTLGEYGRTVVPFSV-- 952
QY 1288 PGDPGRARLLEALLQEAAGLEBPVQHSRHSFSAPEVPAEVAKVLASGNWTPESPSPACQC 1347
DB 953 ---PGTSOLGQQL-----SEH----- 965
QY 1348 SQPGARLLLPDCPAAAGPPPPQAVTSGEVVQNLGRNLSDFLVKTYPRVLRQGLKTKK 1407
DB 966 -----LKDALQAG--QEPREVLG-----DLEEFLLI----- 989
QY 1408 WNEVRYGFSGLGRDPLPSGOELGSRVBEELWALLSPLPGGALDRVLKNLTAWAHSlda 1467
DB 990 FRASVEGGFN-----ERCL--VAASFRDVG 1014
QY 1468 QDSLKTIWNNKWHISWAFVNRASNAILRAHLPPGPBARHAHSITTLNHP-----LNLTK 1522
DB 1015 RTVYNALFNQAYHSPATALAVVDNLLFK--LLCGP--HA-SIVVSNFPQPRSAQAOKD 1069
QY 1523 QISEAALMASSVDVLVISCIVVPFAMSVFASFTLVLTIEERVTRAKHLQLMGSLPTLYWL 1582
DB 1070 QFNEG--RKGFDIALNL--LFMAFLASTFSLASERAVQAKHVQFVSGVHVASFWS 1124
QY 1583 NFLWDMCNLYVPACIVLVILFAPQOQRAYAPANLPALLLLLLLYGWSITPLMYPAFFFFS 1642
DB 1125 ALLWDLISFLIPSLLLLVVFKAPDVRFTRDGHMADTLLLLLLYGWAIPLMYLMNFFFL 1184
QY 1643 VPTAYVVLTCINLFTGINGSMTFVLEL---PSDOKLOEVSRIILKQVELIPHFCLGRG 1699
DB 1185 GAATAYTRLTIFNILSGI-----ATFLMVTIMRIPAVKLEELSKTLDHVLVLNHCIGMA 1240
QY 1700 LIDMVRN-----QAMADAPERLGRDQFQSP--RWEV--VGKNLLAMVIOQ-----P 1742
DB 1241 VSSFYENYETRRYCTSSSEVAHYCKYNIQYQENFYAMSPGVGRFVMAASGCAVLIL 1300
QY 1743 LFLLLTLLLOHRSQLLPQPRVR-----SLPLGEEDEVARERVRVQATQG---D 1791
DB 1301 LFLIETNLLQRLGILCALRRRTLTFLVTRMBVL-PEQDQVADERTRILAPSPDSLLHT 1359
QY 1792 VLVLRLITKVYRQORP--AVDRLCIGIPGECFGLLVNGAGKSTFRVMTGDTLASRG 1849
DB 1360 PLTIKLSKVYE--QRVPELLAVDRSLAVQKGCFLGLGFGAGKTTTFKMLTGEESLTS 1418
QY 1850 EAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLITGREHLELLARLRGVPEAQVATAGS 1909
DB 1419 DAFVGHRISSDVGVKVRQIRIGYCPQFDALLDHTGEMLVYARLARGIPERHIGACVENT 1478
QY 1910 LABRLGSIWADRPAGTYSGNKKKATALALVGDPAVFLDEPTTGMDSARFLNLSLL 1969
DB 1479 LRGLLEPHANKLVRTYSGNKKKSTGIALIGEPAVIFLDEPSTGMDPVARELLMDTVA 1538
QY 1970 AVVREGSVMLTSHSMECEALCSRLAMVNGEFCGLGSPQHLKGRFAAGHTTLTRVPA 2029
DB 1539 RARESGKAIITSHSMECEALCTRLAMVQGFQKLGSPQHLKSPFGSGYSIRAKVQSE 1598
QY 2030 RSOPA---AAFAVAEPFGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAEGHVED 2085

Db 1599 GQCEALEBFKAFVDLTFPGSVLEDEHQGVHYHL-PGRDLSWAKVFGILEKAKERYGVDD 1657
QY 2086 RSVSOTMLEEVFLYFS 2101
Db 1658 YSVSQISLEQVFLSFA 1673

RESULT 14
US-10-000-489-102
; Sequence 102, Application US/10000489
; Patent No. 6794383
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 102
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-489-102

Query Match 3.9%; Score 430; DB 4; Length 162;
Best Local Similarity 54.7%; Pred. No. 4.3e-30;
Matches 76; Conservative 26; Mismatches 37; Indels 0; Gaps 0;
QY 1 MAFWTQLMLLKNFMYRRQRPVQLLVLLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MACWPQLLLWKNLTERRQTCQLLELVANWPLFIFLILSVRLSYPPYEQHECHFPNKA 60
QY 61 LPSAGTVPWLOGLICNVNNTCFPQLTGPEEGRISNFNDLSVRLADARTVLGASAH 120
Db 61 MPSAGTLPWQGLICNANNPCFRYPTEGAEPCGVGNFNKSIVARLFSDDARRLLLYSQKDT 120
QY 121 TLAGLGKLIATLRAARSTA 139
Db 121 SMKMKRVLRITLQOIKKSS 139

RESULT 15
US-09-724-797-8
; Sequence 8, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 328

; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-8
Query Match 3.8%; Score 427; DB 4; Length 328;
Best Local Similarity 34.9%; Pred. No. 2.9e-29;
Matches 112; Conservative 59; Mismatches 142; Indels 8; Gaps 5;
QY 1796 RNLTKVYRGQMPAVDRICLIGIPPGCEFCGLLVGNGAGKTSFRMVVTGDTLASRGAVALG 1855
Db 12 RDLVKVFGQTR--AVDGLDLVVRAGTIHGVLGPGAGKTTAKMLATLMRPTSGTASVIG 69
QY 1856 HSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSGLARLGL 1915
Db 70 HDVVREAAAEVRRRIGLTGQMTSVDMDTGVQNLILAGRLQCLRHASAAARAEQMLEAFDL 129
QY 1916 SWYADRPAGTYSGGNKRLATALALVGDPAVVFLDEPTTGMDPSARRFLWNSLLAVVREG 1975
Db 130 TEVGGRLVKTFSGGQRRRIDVAASMVVTPELLFLDEPTTGIDPRSRSEVWEMIRALVRDG 189
QY 1976 RSVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAG-HTLTLRVPAARSQPA 2034
Db 190 GTVLTTQYLDLADHLADELTLDHGRIVAGQTPPELKASRAAGVLDVRLRDPERRADAG 249
QY 2035 AAFVAABFPGSELREAHGGRRLRFQLPQGRCALARVFGELAVHGAHGEVDFSVSQTMLE 2094
Db 250 ALLAKAVGAAADL-DSDPARLSVRVTDPDRAALA--LGELARAGIH--VDDFTLQGPSLD 304
QY 2095 EVFLYFSKQDQKEDTBEQKE 2115
Db 305 TVFLALTGHSTVDASEEEAE 325

Search completed: December 29, 2004, 22:53:49
Job time : 43.7979 secs

This Page Blank (uspto)

GenCore version 5.1.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 243.611 Seconds
(without alignments)
5068.549 Million cell updates/sec

Title: US-09-995-542-5

Perfect score: 11143

Sequence: 1 MAFWTQMLLLWKQNFMYRR.....QHPKRVSLDDPSTAEVTL 2146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11132	99.9	2146	2 Q9BZC4	Q9BZC4 homo sapien
2	11130	99.9	2146	2 Q81ZY2	Q81ZY2 homo sapien
3	11128	99.9	2146	2 Q9NR73	Q9NR73 homo sapien
4	10249	92.0	2008	2 Q9GS58	Q9GS58 homo sapien
5	8652	77.6	2170	2 Q7TNJ2	Q7TNJ2 rattus norv
6	8600.5	77.2	2159	2 Q91V24	Q91V24 mus musculu
7	5817	52.2	2260	2 Q81UV4	Q81UV4 gallus gall
8	5761.5	51.7	2261	1 ABC1_MOUSE	P41233 mus musculu
9	5751.5	51.6	2261	1 ABC1_HUMAN	Q95477 homo sapien
10	5510.5	49.5	2201	2 Q80ZB2	Q80ZB2 rattus norv
11	5030	45.1	2269	2 Q6T941	Q6T941 canis faml
12	5030	45.1	2269	2 AAR87835	AAR87835 canis fam
13	5029	45.1	2269	2 Q6T940	Q6T940 canis faml
14	5029	45.1	2269	2 AAR87836	AAR87836 canis fam
15	5028	45.1	2269	2 Q6T942	Q6T942 canis faml
16	5028	45.1	2269	2 AAR87834	AAR87834 canis fam
17	4971.5	44.6	2281	2 O02698	O02698 bos taurus
18	4948.5	44.4	2310	1 ABCR_MOUSE	O35600 mus musculu
19	4948.5	44.4	2310	2 AAR57853	AAR57853 mus muscu
20	4936	44.3	2273	1 ABCR_HUMAN	P78363 homo sapien
21	3629.5	32.6	2434	1 ABC2_RAT	Q9eer9 rattus norv
22	3621	32.5	2436	2 Q9HC28	Q9HC28 homo sapien
23	3611	32.4	2436	1 ABC2_HUMAN	Q9BZC7 homo sapien
24	3496.5	31.4	2434	1 ABC2_MOUSE	P41234 mus musculu
25	3437	30.8	1487	2 Q8BPY1	Q8BPY1 mus musculu
26	3351.5	30.1	1771	2 Q76MW7	Q76MW7 homo sapien
27	3351.5	30.1	1771	2 BAA83014	BAA83014 homo sapi
28	2808	25.2	2595	1 AB12_HUMAN	Q86uk0 homo sapien
29	2715	24.4	1416	2 Q6ZP24	Q6ZP24 mus musculu
30	2715	24.4	1416	2 BAC98084	BAC98084 mus muscu
31	2714	24.4	2127	2 Q86W12	Q86W12 homo sapien

32 2701.5 24.2 5058 2 Q86UQ4 Q86UQ4 homo sapien
33 2645.5 23.7 2143 2 Q80T20 Q80T20 mus musculu
34 2503 22.5 1704 2 BAB86781 BAB86781 homo sapi
35 2502 22.5 1704 1 ABC3_HUMAN ABC3_HUMAN
36 2471.5 22.2 1538 2 Q80XT2 Q80XT2 homo sapien
37 2468 22.1 1764 2 Q8MUA3 Q8MUA3 strongyloce
38 2463.5 22.1 1704 1 ABC3_MOUSE ABC3_MOUSE
39 2316.5 20.8 1563 2 Q7PZY9 Q7PZY9 anopheles g
40 2272.5 20.4 1713 2 Q9VRG4 Q9VRG4 anopheles g
41 2184 19.6 1679 2 Q7Q7K8 Q7Q7K8 anopheles g
42 2165.5 19.4 1683 2 Q6XBG3 Q6XBG3 mus musculu
43 2165.5 19.4 1683 2 AAF73044 AAF73044 mus muscu
44 2091.5 18.8 1750 2 Q7QJG5 Q7QJG5 anopheles g
45 2070 18.6 1578 2 Q7Q5K8 Q7Q5K8 anopheles g

ALIGNMENTS

RESULT 1
Q9BZC4
ID Q9BZC4 PRELIMINARY; PRT; 2146 AA.
AC Q9BZC4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ABC transporter member 7.
GN Name=ABCA7;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21328888; PubMed=11435699;
RA Brucardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,
RA Shulenian S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,
RA Mattei M.G., Dean M., Denefle P., Chimini G.;
RT "Comparative analysis of the promoter structure and genomic
RT organization of the human and mouse ABCA7 gene encoding a novel ABCA
RT transporter.";
RL Cytogenet. Cell Genet. 92:264-270 (2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF328787; AAK00959.1; .
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234305 MW; 2391728D5AD97E75 CRC64;

Query Match 99.9%; Score 11132; DB 2; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2144; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAFWTQMLLLWKQNFMYRRQVQLLVLLPFLFFILVAVRHSHHPLEHHECHFPNKP 60
Db 1 MAFWTQMLLLWKQNFMYRRQVQLLVLLPFLFFILVAVRHSHHPLEHHECHFPNKP 60
Qy 61 LPSAGTVPWLOGLICNVNNTCFPOLTPCEEPGRISNFNDSLSRLADARTVLGGASAH 120
Db 61 LPSAGTVPWLOGLICNVNNTCFPOLTPCEEPGRISNFNDSLSRLADARTVLGGASAH 120
Qy 121 TLAGLGLIATLRAARSTAPQPTKQSFPLEPPMLDVAELLTSLRTESLGLALGOAQEPL 180
Db 121 TLAGLGLIATLRAARSTAPQPTKQSFPLEPPMLDVAELLTSLRTESLGLALGOAQEPL 180

Db 121 TLAGLGLIATLRAASTAOPQTKQSPLEPPMLDVAELLTSLRTSGLGALQAOEPL 180
Qy 181 HSLEAAEDLAQELLALRSVLBAILLORPCTSGPLLSLSEALCSVRGSSSTVGPSSLNW 240
Db 181 HSLEAAEDLAQELLALRSVLBAILLORPCTSGPLLSLSEALCSVRGSSSTVGPSSLNW 240
Qy 241 YEASDLMELVQBPESALPOSSLSPACSELIGALDSDHPLSRLLWRLKPLILGKLLFAPD 300
Db 241 YEASDLMELVQBPESALPOSSLSPACSELIGALDSDHPLSRLLWRLKPLILGKLLFAPD 300
Qy 301 TPFTKRLMAQVNRTEBELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLQWODEGRR 360
Db 301 TPFTKRLMAQVNRTEBELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLQWODEGRR 360
Qy 361 QPRPGGRDHMBALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Db 361 QPRPGGRDHMBALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Qy 421 LVSALQALLAHRFPWAGVVFGLGPDSDPTHEPTDLPDGHVRKIRMDIDVTRTNKIR 480
Db 421 LVSALQALLAHRFPWAGVVFGLGPDSDPTHEPTDLPDGHVRKIRMDIDVTRTNKIR 480
Qy 481 DRFDWPGPAADPLTDLRYVMGFFYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
Db 481 DRFDWPGPAADPLTDLRYVMGFFYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
Qy 541 VFLRVLSRSLPLFLTLAWIYSVTLTVKAVVREKTRLDTRWAMGLSRVILWLGWFLSCL 600
Db 541 VFLRVLSRSLPLFLTLAWIYSVTLTVKAVVREKTRLDTRWAMGLSRVILWLGWFLSCL 600
Qy 601 GPFLLSAALLVLKIGDILPYSHPGVVFLLAFAFVATVTSQFLLSAFFSRANLAAACG 660
Db 601 GPFLLSAALLVLKIGDILPYSHPGVVFLLAFAFVATVTSQFLLSAFFSRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVNRDRLPAGGRVAASLLSPVAFGFCESIALLEEQEGEQWENVG 720
Db 661 GLAYFSLYLPVLCVNRDRLPAGGRVAASLLSPVAFGFCESIALLEEQEGEQWENVG 720
Qy 721 TRPTADVLSAQSGLLLDAAALYGLATWYLEAVCPQGYGIPPEWNPFFRRSYWCGRPP 780
Db 721 TRPTADVLSAQSGLLLDAAALYGLATWYLEAVCPQGYGIPPEWNPFFRRSYWCGRPP 780
Qy 781 KSPAPCTPDPKVLVEAPPLSGVSVRSLEKRPFGSPQALRGHSLDFYQGHITAF 840
Db 781 KSPAPCTPDPKVLVEAPPLSGVSVRSLEKRPFGSPQALRGHSLDFYQGHITAF 840
Qy 841 GHNGAGKTTTLLSILSGLFPSPGSAFILGHDVRSMAAIRPHLGVCQYXNVLPDMLTVDE 900
Db 841 GHNGAGKTTTLLSILSGLFPSPGSAFILGHDVRSMAAIRPHLGVCQYXNVLPDMLTVDE 900
Qy 901 HWWFYGRBKGLSAAVGPQDRLLQDVLVSKSVQTRHLSGGMOKLSVAIAFVGSQV 960
Db 901 HWWFYGRBKGLSAAVGPQDRLLQDVLVSKSVQTRHLSGGMOKLSVAIAFVGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLKXVREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLKXVREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTQEKKNQSGSRVGTPOLLA 1080
Db 1021 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTQEKKNQSGSRVGTPOLLA 1080
Qy 1081 LVQHWVFGARLVEBPHLVLVLPYTGANDGSPATLFPRELDTRIAELRLTGYISDTSLE 1140
Db 1081 LVQHWVFGARLVEBPHLVLVLPYTGANDGSPATLFPRELDTRIAELRLTGYISDTSLE 1140
Qy 1141 EIFLKVVEECAADTMEDGSGOHLCTGCIAGLDVTLKMPPOQTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTMEDGSGOHLCTGCIAGLDVTLKMPPOQTALENGEPAGSAPETD 1200
Qy 1201 QGSGFDVAVRGVQWALTRQQLQALLKXFLARSRRLGFAQIVLPALFVGLALVFSLIV 1260
Db 1201 QGSGFDVAVRGVQWALTRQQLQALLKXFLARSRRLGFAQIVLPALFVGLALVFSLIV 1260

Qy 1261 PPRGHYPALRLSPMTYCAQVSFFSEDAPDGPGRARLLLEALLQEAGLEPPVQHSRHFSA 1320
Db 1261 PPRGHYPALRLSPMTYCAQVSFFSEDAPDGPGRARLLLEALLQEAGLEPPVQHSRHFSA 1320
Qy 1321 PEVPAEYAKVLASGNMTPESSPACQSQSGARLLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Db 1321 PEVPAEYAKVLASGNMTPESSPACQSQSGARLLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Qy 1381 NLTCRNLSDFLVKTYPRQGLKTKKVNVEVYGGHSLGGRDPLGSGQELGRSVBELW 1440
Db 1381 NLTCRNLSDFLVKTYPRQGLKTKKVNVEVYGGHSLGGRDPLGSGQELGRSVBELW 1440
Qy 1441 ALLSPILPGGALDRVLKNTAWAHSLLDAQDSIKTWNNKGHSMVAFVNRASNAITLRAHLP 1500
Db 1441 ALLSPILPGGALDRVLKNTAWAHSLLDAQDSIKTWNNKGHSMVAFVNRASNAITLRAHLP 1500
Qy 1501 PGPARHAHSITTLNHPNLTKQOLSEAAALMASSVDVLVSIQVVFAMSFVPASFTVLIEE 1560
Db 1501 PGPARHAHSITTLNHPNLTKQOLSEAAALMASSVDVLVSIQVVFAMSFVPASFTVLIEE 1560
Qy 1561 RVTRAKHLQMLGSLPTLYWGNFLMDMCMNYPACIVVLI FLAFOQRAVYVAPANLPALL 1620
Db 1561 RVTRAKHLQMLGSLPTLYWGNFLMDMCMNYPACIVVLI FLAFOQRAVYVAPANLPALL 1620
Qy 1621 LLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEV 1680
Db 1621 LLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEV 1680
Qy 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLQDGFQSPFLRVEVYVKNLLAMVIQ 1740
Db 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLQDGFQSPFLRVEVYVKNLLAMVIQ 1740
Qy 1741 GPFLFLFTLLQHRSQLLPQPRVRSPLPILGEEDVARERERVVQGTQGDVLVLRNLTK 1800
Db 1741 GPFLFLFTLLQHRSQLLPQPRVRSPLPILGEEDVARERERVVQGTQGDVLVLRNLTK 1800
Qy 1801 VYRGQRPADVRLCIGIPPCGECFLLGVNGAGKTSFTPMVTGDTLASRGAVALAGHSVAR 1860
Db 1801 VYRGQRPADVRLCIGIPPCGECFLLGVNGAGKTSFTPMVTGDTLASRGAVALAGHSVAR 1860
Qy 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSYAD 1920
Db 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSYAD 1920
Qy 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDDPSARRFLWNLSLLAVVREGSRVML 1980
Db 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDDPSARRFLWNLSLLAVVREGSRVML 1980
Qy 1981 TSHSMEECEALCSRLAIVNNGRFRCLGSPHLLGRFAAGHTTLTRVPAARSQPAFAAFVAA 2040
Db 1981 TSHSMEECEALCSRLAIVNNGRFRCLGSPHLLGRFAAGHTTLTRVPAARSQPAFAAFVAA 2040
Qy 2041 EFPQSELRREAHGRLRFLQPPGQRCLARVFGELAVHGAEBHGVDFSVQTMLEEVFLYF 2100
Db 2041 EFPQSELRREAHGRLRFLQPPGQRCLARVFGELAVHGAEBHGVDFSVQTMLEEVFLYF 2100
Qy 2101 SKDQKDEDETEQKAGVGDPAVGLQHPKRVSOQLDDPSTAEVTL 2146
Db 2101 SKDQKDEDETEQKAGVGDPAVGLQHPKRVSOQLDDPSTAEVTL 2146

RESULT 2

Q81ZY2
ID Q81ZY2 PRELIMINARY; PRT; 2146 AA.
AC Q81ZY2;
DT 01-MAR-2003 (T=EMBLrel. 23, Created)
DT 01-MAR-2003 (T=EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T=EMBLrel. 26, Last annotation update)
DE ABC transporter ABCA7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20549028; PubMed=11095984;
RA Kaminski W.E., Piehler A., Schmitz G.;
RT "Genomic organization of the human cholesterol-responsive ABC
RT transporter ABCA7: tandem linkage with the minor histocompatibility
RT antigen HA-1 gene.";
RL Biochem. Biophys. Res. Commun. 278:782-789(2000).
CC - !- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF311060; AAN04657.1; JOINED.
DR EMBL; AF311061; AAN04657.1; JOINED.
DR EMBL; AF311062; AAN04657.1; JOINED.
DR EMBL; AF311063; AAN04657.1; JOINED.
DR EMBL; AF311064; AAN04657.1; JOINED.
DR EMBL; AF311065; AAN04657.1; JOINED.
DR EMBL; AF311066; AAN04657.1; JOINED.
DR EMBL; AF311067; AAN04657.1; JOINED.
DR EMBL; AF311068; AAN04657.1; JOINED.
DR EMBL; AF311069; AAN04657.1; JOINED.
DR EMBL; AF311070; AAN04657.1; JOINED.
DR EMBL; AF311071; AAN04657.1; JOINED.
DR EMBL; AF311072; AAN04657.1; JOINED.
DR EMBL; AF311073; AAN04657.1; JOINED.
DR EMBL; AF311074; AAN04657.1; JOINED.
DR EMBL; AF311075; AAN04657.1; JOINED.
DR EMBL; AF311076; AAN04657.1; JOINED.
DR EMBL; AF311077; AAN04657.1; JOINED.
DR EMBL; AF311078; AAN04657.1; JOINED.
DR EMBL; AF311079; AAN04657.1; JOINED.
DR EMBL; AF311080; AAN04657.1; JOINED.
DR EMBL; AF311081; AAN04657.1; JOINED.
DR EMBL; AF311082; AAN04657.1; JOINED.
DR EMBL; AF311083; AAN04657.1; JOINED.
DR EMBL; AF311084; AAN04657.1; JOINED.
DR EMBL; AF311085; AAN04657.1; JOINED.
DR EMBL; AF311086; AAN04657.1; JOINED.
DR EMBL; AF311087; AAN04657.1; JOINED.
DR EMBL; AF311088; AAN04657.1; JOINED.
DR EMBL; AF311089; AAN04657.1; JOINED.
DR EMBL; AF311090; AAN04657.1; JOINED.
DR EMBL; AF311091; AAN04657.1; JOINED.
DR EMBL; AF311092; AAN04657.1; JOINED.
DR EMBL; AF311093; AAN04657.1; JOINED.
DR EMBL; AF311094; AAN04657.1; JOINED.
DR EMBL; AF311095; AAN04657.1; JOINED.
DR EMBL; AF311096; AAN04657.1; JOINED.
DR EMBL; AF311097; AAN04657.1; JOINED.
DR EMBL; AF311098; AAN04657.1; JOINED.
DR EMBL; AF311099; AAN04657.1; JOINED.
DR EMBL; AF311100; AAN04657.1; JOINED.
DR EMBL; AF311101; AAN04657.1; JOINED.
DR Genew; HGNC:37; ABCA7.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234421 MW; 33A128082D7B5BAF CRC64;

Query Match

99.9%; Score 11130; DB 2; Length 2146;

Db 1021 SPLFLRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEBKKSGQSRVCTPQLLA 1080
Qy 1081 LVQHWPFGARLVEELPHELVLVLVLYTGAHDSFATLFRDLRLAELRLTGYGISDTSLE 1140
Db 1081 LVQHWPFGARLVEELPHELVLVLVLYTGAHDSFATLFRDLRLAELRLTGYGISDTSLE 1140
Qy 1141 EIFLKVVEECAADTDMEDSGCGHLCYAGLDVTLRLKMPQSTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTDMEDSGCGHLCYAGLDVTLRLKMPQSTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLIARSRRGLFAQIVLPALFVGLALVFSILIV 1260
Db 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLIARSRRGLFAQIVLPALFVGLALVFSILIV 1260
Qy 1261 PPFCHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGLEPPVQVHSHRFS 1320
Db 1261 PPFCHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGLEPPVQVHSHRFS 1320
Qy 1321 PEVPAEVAKVLASGNWTPESPSPACQSCOPGARELLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Db 1321 PEVPAEVAKVLASGNWTPESPSPACQSCOPGARELLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Qy 1381 NLTCRNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPLSGOELGRSVEELW 1440
Db 1381 NLTCRNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPLSGOELGRSVEELW 1440
Qy 1441 ALLSPLPGGALDRVLKNTAWAHSALDAQDSLKIFWNNKWHSMVAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSALDAQDSLKIFWNNKWHSMVAFVNRASNAILRAHLP 1500
Qy 1501 PGPARHAHSITTLNHLNPLNKEQLSEALMASSVDVLVSVICVWFAMSFVPASFTVLVIEE 1560
Db 1501 PGPARHAHSITTLNHLNPLNKEQLSEALMASSVDVLVSVICVWFAMSFVPASFTVLVIEE 1560
Qy 1561 RVTRAKHLQMGGLSPTLYLWGNFLWDMCNVLPACIWLIFLAFQORAVVAPANLPA 1620
Db 1561 RVTRAKHLQMGGLSPTLYLWGNFLWDMCNVLPACIWLIFLAFQORAVVAPANLPA 1620
Qy 1621 LLLLYLWGSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQKLOEV 1680
Db 1621 LLLLYLWGSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQKLOEV 1680
Qy 1681 SRLIKQVFLIPPHFCICGLIDMVRNQAMADAFERLGDQFQSPRLRVEWVGKLLAMV 1740
Db 1681 SRLIKQVFLIPPHFCICGLIDMVRNQAMADAFERLGDQFQSPRLRVEWVGKLLAMV 1740
Qy 1741 GPLFLLFTLLQHRSQLLPQVRVRSPLLGEEDEVDARERVVQAGTQGDVLRNLTK 1800
Db 1741 GPLFLLFTLLQHRSQLLPQVRVRSPLLGEEDEVDARERVVQAGTQGDVLRNLTK 1800
Qy 1801 VYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFTFMVGTDLASRGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFTFMVGTDLASRGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAGTAGSLARGLSNYAD 1920
Db 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAGTAGSLARGLSNYAD 1920
Qy 1921 RPAGTYSGNKRKLATALLVGDPAVFLDPTTGMPPSARRFLWNSLLAVVREGSRVWL 1980
Db 1921 RPAGTYSGNKRKLATALLVGDPAVFLDPTTGMPPSARRFLWNSLLAVVREGSRVWL 1980
Qy 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAA 2040
Db 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAA 2040
Qy 2041 BFPSCSEALREAGGBRLRQLPPGRCALARVGEALVHGAEGVDFSVQTMLEEVLYF 2100
Db 2041 BFPSCSEALREAGGBRLRQLPPGRCALARVGEALVHGAEGVDFSVQTMLEEVLYF 2100
Qy 2101 SKDQKQEDTEQKEAGVGDPAFGLQHPKRVVSQFLLDDPSTAEIVL 2146
Db 2101 SKDQKQEDTEQKEAGVGDPAFGLQHPKRVVSQFLLDDPSTAEIVL 2146

RESULT 3

Q9NR73
ID Q9NR73 PRELIMINARY; PRT; 2146 AA.
AC Q9NR73;
DT 01-OCT-2000 (T-EMBurel. 15, Created)
DT 01-OCT-2000 (T-EMBurel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBurel. 26, Last annotation update)
DE Macrophage ABC transporter.
GN Name=ABCA7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20334305; PubMed=10873640;
RA Kaminski W.E., Orso E., Diederich W., Klucken J., Drobnik W.,
RA Schmitz G.;
RT Identification of a novel human sterol-sensitive ATP-binding cassette
transporter (ABCA7).
RL Biochem. Biophys. Res. Commun. 273:532-538 (2000).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF250238; AAF85794.1; -.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005524; F:ATP binding; TAS.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; TAS.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF000005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234468 MW; 679B16EB2D75FF0D CRC64;
Query Match 99.9%; Score 11128; DB 2; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAFTQMLLLKNNMYRRQPVQLLVLLVPLFLPILVAVRSHSPPLEHHECHFPNKP 60
Db 1 MAFTQMLLLKNNMYRRQPVQLLVLLVPLFLPILVAVRSHSPPLEHHECHFPNKP 60
Qy 61 LPSAGTVPWLQGLICNVNNTCFPQLTGEPEPGRLSNFNDSLVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPWLQGLICNVNNTCFPQLTGEPEPGRLSNFNDSLVSRLLADARTVLGGASAH 120
Qy 121 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVALLTSLRTESLGLALQAOEPL 180
Db 121 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVALLTSLRTESLGLALQAOEPL 180
Qy 181 HSLLAEADLAOELLARLSVELRALLQRPRTSGPPELLELSEALCSVRGSSSTVGPSSLNW 240
Db 181 HSLLAEADLAOELLARLSVELRALLQRPRTSGPPELLELSEALCSVRGSSSTVGPSSLNW 240
Qy 241 YEASDLMELVQGPESALPDSSLSLPACSELIGALDSHPLSRLLWRLKPLILGLKLLFAPD 300
Db 241 YEASDLMELVQGPESALPDSSLSLPACSELIGALDSHPLSRLLWRLKPLILGLKLLFAPD 300
Qy 301 TPFTKRLMAOVNRTFEBTLTLRDVREVMELGPRIFTFMNDSSNVAMQLRLLQMDGRR 360
Db 301 TPFTKRLMAOVNRTFEBTLTLRDVREVMELGPRIFTFMNDSSNVAMQLRLLQMDGRR 360
Qy 361 QPFGGGRDHMEALRSFLDPCGGYSGWODAHADVGHVGLTGRVTECLSLDKLEAPSEAA 420
Db 361 QPFGGGRDHMEALRSFLDPCGGYSGWODAHADVGHVGLTGRVTECLSLDKLEAPSEAA 420
Qy 421 LVSRALQLLAHFRFWAGVWFLGPEDSSDPTHEHTPDILGPGHVRKIRMDIDVVTNRKIR 480
Db 421 LVSRALQLLAHFRFWAGVWFLGPEDSSDPTHEHTPDILGPGHVRKIRMDIDVVTNRKIR 480

QY 481 DRFPDGPAAADPLTDLRVYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPPCYVD 540
DB 481 DRFPDGPAAADPLTDLRVYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPPCYVD 540
QY 541 VFLRVLSRSLPLFLTLTAWIYSVTLVKAVVREKETRLDTRAMAGLSRAVLWLGWFLSCL 600
DB 541 VFLRVLSRSLPLFLTLTAWIYSVTLVKAVVREKETRLDTRAMAGLSRAVLWLGWFLSCL 600
QY 601 GFPLLSAALLVLVLKGDILPYSHPGVWFLFLAAFAVATVTFOSFLLSAFFSFRANLAAACG 660
DB 601 GFPLLSAALLVLVLKGDILPYSHPGVWFLFLAAFAVATVTFOSFLLSAFFSFRANLAAACG 660
QY 661 GLAYFSLYLPYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGGAQWHNVG 720
DB 661 GLAYFSLYLPYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGGAQWHNVG 720
QY 721 TRPTADVFLSLAQVSGLLLDLDAALYGLATWYLEAVCPGQYGIPEPWNFPFRSRVWCGPRPP 780
DB 721 TRPTADVFLSLAQVSGLLLDLDAALYGLATWYLEAVCPGQYGIPEPWNFPFRSRVWCGPRPP 780
QY 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLSLDFYQGHITAPL 840
DB 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLSLDFYQGHITAPL 840
QY 841 GHNGAGKTTTSLISGLPFPSPGSAFILGHVDRSSMAAIRPHLGVCPQVNVLFDMLTVD 900
DB 841 GHNGAGKTTTSLISGLPFPSPGSAFILGHVDRSSMAAIRPHLGVCPQVNVLFDMLTVD 900
QY 901 HWVFYGRUKGLSAAVGPQDRLQDVLVSKQSVQTRHLSGGMQRKLSVAIFVGGSQV 960
DB 901 HWVFYGRUKGLSAAVGPQDRLQDVLVSKQSVQTRHLSGGMQRKLSVAIFVGGSQV 960
QY 961 VILDEPTAGVDPASRGITWELLKYREGTTLTSTHLLDEAELLGDRVAVAGRLCCG 1020
DB 961 VILDEPTAGVDPASRGITWELLKYREGTTLTSTHLLDEAELLGDRVAVAGRLCCG 1020
QY 1021 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKKNGSQSGRVGTPQLIA 1080
DB 1021 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKKNGSQSGRVGTPQLIA 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGANDGFSATLFRDLTRLAELRLTGYISDTSLE 1140
DB 1081 LVQHWVPGARLVEELPHELVLVLPYTGANDGFSATLFRDLTRLAELRLTGYISDTSLE 1140
QY 1141 EFLKVVBECAADTDMEGSCGQHLCGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
DB 1141 EFLKVVBECAADTDMEGSCGQHLCGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAGRVQGWALTROQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFSLIV 1260
DB 1201 QGSGPDVAGRVQGWALTROQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFSLIV 1260
QY 1261 PPFHYPALRLSPTWYGAQVSPFSDAPGDPGRARLLEALLQEAELPVPVQSHSRFSA 1320
DB 1261 PPFHYPALRLSPTWYGAQVSPFSDAPGDPGRARLLEALLQEAELPVPVQSHSRFSA 1320
QY 1321 PEVPAEVAKLASGNWTPESPACQSGOPGARRLLPDCPAAAGPPPPQAVTGSGEVYQ 1380
DB 1321 PEVPAEVAKLASGNWTPESPACQSGOPGARRLLPDCPAAAGPPPPQAVTGSGEVYQ 1380
QY 1381 NLTRGNLSDFLVKTYPRLVROGLTKTKWVNEVRYGSGFLGGRDPLGSGQELGRSVEELW 1440
DB 1381 NLTRGNLSDFLVKTYPRLVROGLTKTKWVNEVRYGSGFLGGRDPLGSGQELGRSVEELW 1440
QY 1441 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWNKGNWSWAFVNRASNAILRAHLP 1500
DB 1441 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWNKGNWSWAFVNRASNAILRAHLP 1500
QY 1501 PGPARRHAHSITTLNHPNLTKQLSEALMASSVDVLVSVVFPAMSVFASFTLVLEEE 1560
DB 1501 PGPARRHAHSITTLNHPNLTKQLFEALMASSVDVLVSVVFPAMSVFASFTLVLEEE 1560

QY 1561 RVTRAKHLQMGGLSPTLYWLGNFMDMCNLYLVPACIVLVILFLAFOORAVYAPANLPALL 1620
DB 1561 RVTRAKHLQMGGLSPTLYWLGNFMDMCNLYLVPACIVLVILFLAFOORAVYAPANLPALL 1620
QY 1621 LLLLYLWGSITPLMYPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLELSDQKLQEV 1680
DB 1621 LLLLYLWGSITPLMYPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLELSDQKLQEV 1680
QY 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGDGFQFQSPLEWVGKLLAMVIO 1740
DB 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGDGFQFQSPLEWVGKLLAMVIO 1740
QY 1741 GFLFLFTLLLOHRSQLLPQPRVRSPLLIGEEDVARERERVVQCATQGDVLVLRNLTK 1800
DB 1741 GFLFLFTLLLOHRSQLLPQPRVRSPLLIGEEDVARERERVVQCATQGDVLVLRNLTK 1800
QY 1801 VTRGQMPADVRLCLGIPGECFGLLVGNVAGKTSFTRMVTGDTLASRGEAVLAGHSVAR 1860
DB 1801 VTRGQMPADVRLCLGIPGECFGLLVGNVAGKTSFTRMVTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLSMGYCPQSDAIFELTGREHLELLARLRCVPEAOVATAGSLARLGLSWYAD 1920
DB 1861 EPSAAHLSMGYCPQSDAIFELTGREHLELLARLRCVPEAOVATAGSLARLGLSWYAD 1920
QY 1921 RPAGTYSNGNKKLATALALVGDPAVVFDEPTTGMDPSARRFLMNSLLAVVREGRSVML 1980
DB 1921 RPAGTYSNGNKKLATALALVGDPAVVFDEPTTGMDPSARRFLMNSLLAVVREGRSVML 1980
QY 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLTLRVPAARSQPAAFVAA 2040
DB 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLTLRVPAARSQPAAFVAA 2040
QY 2041 EPPGSELREAGRLRFOLPPGRCALARVFCGLAVHGAHEGVDESVQTMLEEVLYF 2100
DB 2041 EPPGSELREAGRLRFOLPPGRCALARVFCGLAVHGAHEGVDESVQTMLEEVLYF 2100
QY 2101 SKDQKQSDTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146
DB 2101 SKDQKQSDTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 4

Q96S58 PRELIMINARY; PRT; 2008 AA.
AC Q96S58;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ABCA-SSN.
GN Name=ABCA7/ABCA-SSN;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21255283; PubMed=11355874;
RA Tanaka A., Ikeda Y., Abe-Dohmae S., Arakawa R., Sadanami K.,
RA Kidera A., Nakagawa S., Nagase T., Aoki R., Kioka N., Amachi T.,
RA Yokoyama S., Ueda K.;
RT "Human ABCA1 Contains a Large Amino-Terminal Extracellular Domain
Homologous to an Epitope of Sjogren's Syndrome";
RL Biochem. Biophys. Res. Commun. 283:1019-1025(2001)
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB055390; BAB62294.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.


```
Qy 1569 QLMGGLSPTLYWLNFLWDMCNVLPACIVVLIIFLAFQORAYVAPANLPALLILLILLYGW 1628
Dy 1594 QLVSGLPQTLWLNFLWDMCNVLYVAVICVVLIFLAFQKAYVAPENLPALLILLILLYGW 1653
Qy 1629 SITLMPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLELPSDOKLQEVSRILKQVF 1688
Dy 1654 SITLMPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLELPSDOKLQEVSRILKQVF 1713
Qy 1689 LIFPHFCGLGRGLDMVRNQAMADAFERLGDQFQSPLRWEVVGKLLAMVIGQFLFLFT 1748
Dy 1714 LIFPHFCGLGRGLDMVRNQAMADAFERLGDQFQSPLRWDIIGKLLAMVAQGFLLIT 1773
Qy 1749 LLLQHRSQLPQPRVRSPLPILGEDEDEVARERERVQATGQDVILVRLNLTKVYRGQMP 1808
Dy 1774 LLLQHRNLLPQPKSRLLPPLIGEDEDVVRERERVTKATGQDVILVRLDUTKVYRGQSP 1833
Qy 1809 AVDRCLGIPGECFGLLLGVNGAGKTSFRVMVTGDTILASRGEAVLAGHSVAREPSAAHLS 1868
Dy 1834 AVDRCLGIPGECFGLLLGVNGAGKTSFRVMVTGDTILPSSGEAVLAGHVAQEPSAAHRS 1893
Qy 1869 MGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLSWYADRPAGTYSG 1928
Dy 1894 MGYCPQSDAIFDLTGREHLELFARLGVPEAQVAQTAGSGLVRLGLPSYADRPAGTYSG 1953
Qy 1929 GNKEKLATALVGDPAVFLDEPTTGMDPSAREFLNLSLLAVVREGSVMLTSHSMEEC 1988
Dy 1954 GNKEKLATALVGDPAVFLDEPTTGMDPSAREFLNLSLLVVRREGSVMLTSHSMEEC 2013
Qy 1989 EALCSRLAIVMNGFRCLGSPHQLKGRFAAGHTLTLRVPAARSQPAAFVAAEPGSELR 2048
Dy 2014 EALCTRLAIVMNGFRCLGSAHLKSRFGAGHTLTLRVPPDQPEATAFIVTFPPDAELR 2073
Qy 2049 EAHGGRRLRFOLPPGRCRCALARVFGELAVHGAHGVEDFSVSQTMLEEVLYFSKQCKDE 2108
Dy 2074 EVHGSRLRFOLPPGCGCTLARVPFELAAQKAHGVEDFSVSQTMLEEVLYFSKQCKDE 2133
Qy 2109 DTEQKEAGVGVDPAPGLQHPKRVSOFLDDPSTAEVTL 2146
Dy 2134 EGSGQETETREVS-TPGLQHPKRVSRFLEDPSSVETVI 2170

RESULT 6
Q91V24 PRELIMINARY; PRT; 2159 AA.
AC Q91V24;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ATP-binding cassette transporter sub-family A member 7.
GN Name=Abca7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=O1129, and DBA/2;
RX MEDLINE=21328886; PubMed=11435699;
RA Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,
RA Shulerin S., Arnold I., Naudin L., Lafargue C., Rosier M., Jordan B.,
RA Mattei M.G., Dean M., Densfle P., Chimini G.;
RT "Comparative analysis of the promoter structure and genomic
RT organization of the human and mouse ABCA7 gene encoding a novel ABCA
RT transporter.";
RL Cytochrome. Cell Genet. 92:264-270(2001).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF287142; AAK56863.1; -.
DR EMBL; AF287141; AAK56862.1; -.
DR HSP; P58301; 1F2U.
DR MGI; MGI:1351646; Abca7.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
```

```
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR02016; Peroxidase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 2159 AA; 236081 MW; CD2B3FE0D8B822B CRC64;

Query Match 77.2%; Score 8600.5; DB 2; Length 2159;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 1676; Conservative 162; Mismatches 280; Indels 69; Gaps 11;

Qy 1 MAFWTQMLLLWGNFMYRRRQPVQLLVELLWPLFLFILLVAVRSHHPLEHHECHFPNKP 60
Dy 1 MALGTQMLLLWKNYTYRRRQPIQLLVELLWPLFLFILLVAVRSHHPLEHHECHFPNKP 60
Qy 61 LPSAGTVPMLQGLICNVNNTCFPLQTPGEBPGRLSNFNDLSVRLADARTVLGGASAH 120
Dy 61 LPSAGTVPMLQGLICNVNNSCFQHPPTFGKPGVLSNFKDSLIRLLADRTVLGGHSIQD 120
Qy 121 TLAGLGLKIATLRAASTAQ-----OPTKQSPLEPPMLDVAELLTSLRTESTLALGQA 176
Dy 121 MLDALGKLIPLVRAVGGARPQSDQPTSG-----SVTKLEKILQASLDPLVGLQA 173
Qy 177 QEPHLSLEAAEDLAQELLALRSILVELRALLQPRGTSGPLELLSEALCSVRGSSIVGP 236
Dy 174 QDSMRKFSDAIRDLAQELLTLPSLMELRALLRPRGSAGSLELVSEALCSVKPSSPGGL 233
Qy 237 SLNMYEASDLMELVGQEPESALPDSLSLSPACSELIGALDSHPLSRLRLKPLILKLL 296
Dy 234 SLNMYEANQLNEFMGFEVAPALPDNLSLSPACSEFVGTLDHDPVSRLLWRLKPLILKIL 293
Qy 297 FAPDTPPTFKLMAQVARTFEELTLRLDRVREVMGLPRIFTFMVDNSNVAMLQRLQMD 356
Dy 294 FAPDTNFTKLMQVNTQPEELALLRDLHMLWGLVGFQIFNFMNDSTNVAMLQRLLDVGG 353
Qy 357 EGRQRPGRGDHWEALRSFLDPGSGGYSQDHAADVGLVGLTRVTECLSIDKLEAAP 416
Dy 354 TGORQQTTPRAQKKEAIKDFLPSRGGYSWREAHDMGLAGILGQMCECVSLDKLEAVP 413
Qy 417 SEALVSRALQLLAERFWAGVVFGLGPESSDPTTEHTPDLPGLGHVRIKIRMDIDVVTRT 476
Dy 414 SEALVSRALLELGERRLMAGIVFLSPHEPLDPSSELSPALSPGLHLPFKIRMDIDVVTRT 473
Qy 477 NKIRDRFMDPGPAADPLTLRLYVVGGFVYLQDLIVERAAVRLVSGANPRAGLYLQOMPYC 536
Dy 474 NKIRDRFMDPGSADPFMDLRYVVGGFVYLQDLLEQAAVRLVGGNSRTGLYLQOMHPHC 533
Qy 537 YVDDVFLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLRDMRANGLSRAVLWLGWF 596
Dy 534 YVDDVFLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLRDMRANGLSRAVLWLGWF 593
Qy 597 LSCGLGPFLLSAALLVLVLGDLILPYSHPGVWFLFAAFVAVATVTSQFLLSAFPSRANLA 656
Dy 594 LSCGLGPFLLSAALLVLVLGDLILPYSHPGVWFLFAAFVAVATVTSQFLLSAFPSRANLA 653
Qy 657 AACGGLAYFSLYLPYVLCVAVMRDLRPPAGGRVAAASLLSPVAFGCGCSLALLEGGEGAQW 716
Dy 654 AACGGLAYFALYLPYVLCVAVMRDLRPPAGGRVAAASLLSPVAFGCGCSLALLEGGEGAQW 713
Qy 717 HNVGTRPTADVLSLAQVSGLLILLDAALYGLATWYLEAVCPGOYGIPEPMNFPERRSYWCG 776
Dy 714 HNLGTGPAEDVFSLAQVSAFLLLDADVIYGLALWYLEAVCPGOYGIPEPMNFPERRSYWCG 773
Qy 777 PRPKGSPACPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPGPSQPQALRGLSLDFYOGHI 836
```

Db 774 PGPSSVLAAPQDPKVLVEBPPIGLVPGVSIIRGLKHFRCPCOPALQGLNLDYEGHI 833
Qy 837 TAFLEHNGAGKTTTILSILSGFPPPGGSAFIFLGHVDRSSMAAIRPHLGVCPQYNVLFDM 896
Db 834 TAFLEHNGAGKTTTILSILSGFPPPGGSSASILGHVDQTNMAAIRPHLGI CQYNVLFDM 893
Qy 897 TVDEHVVYGRILKGLSAAVWGPEDQRLLODVLGSKSVQTRHLSGGHOKLSVAIAFVG 956
Db 894 TVEEHVVYGRILKGLSAAVWGPEDQRLLODVLGSKSVQTRHLSGGHOKLSVAIAFVG 953
Qy 957 GSOVILDEPTAGVDPASRRGIWELLKXVREGRTILSTHLDDEALLGDRVAVAGRL 1016
Db 954 GSRVIMDEPTAGVDPASRRGIWELLKXVREGRTILSTHLDDEALLGDRVAVAGRL 1013
Qy 1017 CCCGFLFLRRHGLGYVLTIVKARLPLTTNE-KADTMEGSDVTRQKNGSQSGS---- 1071
Db 1014 CCCGFLFLRRHGLGYVLTIVKSSQSLVTHDAKGDSE-----DPRREKSDNGRTSDT 1068
Qy 1072 -----RVGTPQLLALVQHVWPGARIVVEELPHLVLVLPVYGAHD 1110
Db 1069 APTRGTSKSNQAPAPGAVPTPSTARILELVQHVPGQALVEDLPHELLVLPVYAGALD 1128
Qy 1111 GSPATLFLRLTRLAELRTGVISDTLSLEIFLKVVECAADTDMEDGSCGHLCTGIA 1170
Db 1129 GSFAMVFOELDQOELLGLTGVISDTNLLEEIFLKVVED--AHREGGDSRPQLHLRT--- 1183
Qy 1171 GLDVTLRKMPQOETALENGEPAGSAPETDQSGPDVGRVQGWALTRQOLQALMLKRL 1230
Db 1184 ---CTPQPTGPEASVLENGELA-----PQGLAPNA-AQVQGWTLTQQLRALLHKRFL 1233
Qy 1231 LARRSRRLGFAQIVLPALFVGLALVSLIVPPFGHYPALRISPTMYGAQVSFFSDEAPGD 1290
Db 1234 LARRSRRLGFAQIVLPALFVGLALVSLIVPPFGHYPALRISPTMYGAQVSFFSDEAPGD 1293
Qy 1291 PGRARLLALQOAGLEBPVQVHSHR-----FSAPEDVPAEVAKVILASGNWTP 1339
Db 1294 PNRMKLLALLGEAGLQEPMSQDKDARGSECTHSLACYFTVPEVDPVASILASGNWTP 1353
Qy 1340 SPSPACQCSQPCARLLPDCPAAAGPPPPQAVTSGVGEVQNLTRGNLSDFLVKTYPRLV 1399
Db 1354 SPSPACQCSQPCARLLPDCPAGAGPPPPQAVAGLGEVQNLTRGNLSDFLVKTYPRLV 1413
Qy 1400 RQGLTKKWNVEVRYGGFSLGGRDPGLPSGOBLGRSVBELWALLSPLPGGALDRVLKMLT 1459
Db 1414 RQGLTKKWNVEVRYGGFSLGGRDPGLPSGOBLGRSVBELWALLSPLPGGALDRVLKMLT 1473
Qy 1460 AWASLDAQDSLIKIFNNKGMHSMVAFVNRASNAILRAHPLPGPARHAHSITTLNHLNL 1519
Db 1474 QWALGLDARNLSLIKIFNNKGMHSMVAFVNRANGLLHALLPSGPPVRAHHSITTLNHLNL 1533
Qy 1520 TKEQLSEALMASSVDVLVSI CVVFPAMSFVPASFTLVLEERVTREAKHLQMLGGLSPTLY 1579
Db 1534 TKEQLSEALMASSVDVLVSI CVVFPAMSFVPASFTLVLEERVTREAKHLQMLGGLSPTLY 1593
Qy 1580 WLGNFLWDMCNVLPACIVVILFLAQORAYVAPANLPALELLLLLYGWSITPLMYPASF 1639
Db 1594 WLGNFLWDMCNVLPACIVVILFLAQORAYVAPANLPALELLLLLYGWSITPLMYPASF 1653
Qy 1640 FFSVSTAYVVLTCINLFIGINGSMATVLELFSQKLOEVSRIKLVFLIPHPCLGRG 1699
Db 1654 FFSVSTAYVVLTCINLFIGINGSMATVLELFSQKLOEVSRIKLVFLIPHPCLGRG 1713
Qy 1700 LIDMVRNQMADAFERLGRQFQSPLEVEVCKNLLAMVIOGPLELLTLLLOHRSQLLP 1759
Db 1714 LIDMVRNQMADAFERLGRQFQSPLEVEVCKNLLAMVIOGPLELLTLLLOHRSQLLP 1773
Qy 1760 OPRVRSPLLGEDEEDVARERERVQCATQGDVLVRLNLTIKYVQRQMPADVRLCLGIPP 1819
Db 1774 QSKPRLPLLGEDEEDVARERERVQCATQGDVLVRLNLTIKYVQRQMPADVRLCLGIPP 1833
Qy 1820 GECFGLLVNGAGKTTSTFRMTVGTDLASRGEAVLAGHSVAREPSAAHLSMGYCPSQDAIF 1879
Db 1834 GECFGLLVNGAGKTTSTFRMTVGTDLASRGEAVLAGHSVAREPSAAHLSMGYCPSQDAIF 1893

Qy 1880 ELLTGREHLELLARLGRVPEAQVTAQTAGSLARLGLSWYADRPACTYSGGNKRKLATALA 1939
Db 1894 DLLTGREHLELLARLGRVPEAQVTAQTAGSLARLGLSWYADRPACTYSGGNKRKLATALA 1953
Qy 1940 LVGDPVAVFLDEPTTGMDDPSARRFLWNSLLAVVREGRSVMLTSHSMEECEALCSRLAIMV 1999
Db 1954 LVGDPVAVFLDEPTTGMDDPSARRFLWNSLLAVVREGRSVMLTSHSMEECEALCSRLAIMV 2013
Qy 2000 NGRFRCILGSPHLLKGRFAAGHTLTLRVPAAQSQAAPAAAFVAAEFPGSELREAHGGRRLRQ 2059
Db 2014 NGRFRCILGSSQHLKGRFGAGHTLTLRVPDOPFPAIAFIRITFFGAEIREVHGSRLRQ 2073
Qy 2060 PPGRCALARYFGELAVHGAEGHVEDFSVQTMLEEVFLYFSKQGDDETEEQKEAGVG 2119
Db 2074 PPGRCCTLTRVRELAAQGRAHGVDFSVQTMLEEVFLYFSKQGDDETEEQKEAGVG 2132
Qy 2120 VDPAGLOHPKRVSQFLDDPSTAEVTL 2146
Db 2133 EVSKPGRQHPKRVSRFLEDPSSVETMI 2159

RESULT 7

Q8UVV4 PRELIMINARY; PRT; 2260 AA.
AC Q8UVV4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATP-binding cassette transporter 1.
GN Name=ABCA1;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Attie A.D., Brooks-Wilson A.R., Walker D., McManus B.,
RA Gray-Keller M.P., MacDonald M.L.E., Roomp K., Tebon A., Zhang L.-H.,
RA Mulligan J., Senses C., Bitgood J.J., Cook M.E., Kastelein J.J.P.,
RA Hayden M.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF362377; AAL56247.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding.
KW ATP-binding.
SQ SEQUENCE 2260 AA; 254070 MW; 19D137F342F98662 CRC64;
Query Match 52.2%; Score 5817; DB 2; Length 2260;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1151; Conservative 369; Mismatches 594; Indels 174; Gaps 25;
Qy 1 MAFWTQLMLLWKNFYRRRRQPVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60
Db 1 MAFWTQLMLLWKNFYRRRRQPVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKA 60
Qy 61 LPSAGTVPWLOGLCNVNTCFPQLTGCEERGLSNFNDLSVRLADATVILGASAH 120
Db 61 MPSAGTLPWIOGLIICNANNPCFRTPTGSGPVGNFASIVSRFLSDAKRLLYSQDT 120
Qy 121 TLALGLKLIATLRAA----- 135

RESULT 8
ID ABC1 MOUSE STANDARD; PRT; 2261 AA.
AC P41233;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette transporter 1) (ATP-binding cassette 1) (ABC-1).
GN Name=Abc1; Synonyms=Abcl1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Macrophage;
RX MEDLINE=94375008; PubMed=8088782;
RA Luciani M.F., Denizot F., Savary S., Mattei M.-G., Chimini G.;
RT "Cloning of two novel ABC transporters mapping on human chromosome 9".
RT Genomics 21:150-159 (1994).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J;
RC MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RA Qiu Y., Cavellier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies revealing novel regulatory sequences".
RT Genomics 73:66-76 (2001).
RL [3]
RN INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not ABCG1 via an LXR-independent pathway".
RT J. Lipid Res. 43:952-959 (2002).
RL [4]
RN DOWN-REGULATION BY ENDOTOXIN.
RX PubMed=12777468; DOI=10.1194/jlr.M300100-JLR200;
RA Khovidhunkit W., Moser A.H., Shigenaga J.K., Grunfeld C., Feingold K.R.;
RT "Endotoxin down-regulates ABCG5 and ABCG8 in mouse liver and ABCA1 and ABCG1 in J774 murine macrophages: differential role of LXR".
RL J. Lipid Res. 44:1728-1736 (2003).
CC -!- FUNCTION: CAMP-dependent and sulfonyleurea-sensitive anion transporter. Key gatekeeper influencing intracellular cholesterol transport (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed in adult tissues. Highest levels are found in pregnant uterus and uterus.
CC -!- INDUCTION: Down-regulated by endotoxins (LPS) or cytokines (TNF and IL-1) in J774 macrophages. The down-regulation by endotoxin in macrophages is not likely to be mediated by the liver X receptor/retinoic X receptor (LXR/RXR).
CC -!- DOMAIN: Multifunctional polypeptide with two homologous halves, each containing an hydrophobic membrane-anchoring domain and an ATP binding cassette (ABC) domain.
CC -!- PTM: Phosphorylation on Ser-2054 regulates phospholipid efflux (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; X75926; CAA53530.1; ALT_INIT.
DR EMBL; AF287263; AAG39073.1; ALT_INIT.

DR MGD; MGI:99607; Abcal.
DR GO; GO:0008203; P:cholesterol metabolism; IDA.
DR GO; GO:0030301; P:cholesterol transport; IDA.
DR GO; GO:0042158; P:lipoprotein biosynthesis; IMP.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD00006; ABC_transporter; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Phosphorylation; Transmembrane; Transport.
FT TRANSMEM 26 42 Potential.
FT TRANSMEM 640 656 Potential.
FT TRANSMEM 690 706 Potential.
FT TRANSMEM 717 733 Potential.
FT TRANSMEM 749 765 Potential.
FT TRANSMEM 771 787 Potential.
FT TRANSMEM 1041 1057 Potential.
FT TRANSMEM 1351 1367 Potential.
FT TRANSMEM 1661 1677 Potential.
FT TRANSMEM 1708 1724 Potential.
FT TRANSMEM 1737 1753 Potential.
FT TRANSMEM 1775 1791 Potential.
FT TRANSMEM 1854 1870 Potential.
FT NP_BIND 933 940 ATP (Potential).
FT NP_BIND 1946 1953 ATP (Potential).
FT MOD_RES 1042 1042 Phosphoserine (by PKA) (By similarity).
FT MOD_RES 2054 2054 Phosphoserine (by PKA) (By similarity).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 98 98 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 161 161 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 244 244 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 292 292 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 337 337 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 349 349 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 478 478 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 489 489 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 820 820 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1144 1144 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1294 1294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1453 1453 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1499 1499 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1504 1504 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1637 1637 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2044 2044 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2238 2238 N-linked (GlcNAc...) (Potential).
FT CONFLICT 1567 1568 Missing (in Ref. 2).
FT CONFLICT 2024 2024 Missing (in Ref. 2).
SQ SEQUENCE 2261 AA; 253972 MW; F0C3C5F1CFE34F1 CRC64;

Query Match 51.7%; Score 5761.5; DB 1; Length 2261;
Best Local Similarity 49.9%; Pred. No. 0;
Matches 1147; Conservative 370; Mismatches 587; Indels 195; Gaps 27;
Qy 1 MAFWTQLMLLWKNMYRRQPVOLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60
Db 1 MACPQLRLLLWKNLTFRRTQCLLEVAWPLFILILSVRLSYPPYEQEHCFFPNKA 60
Qy 61 LPSAGTVFWLGLICNVNNTCPQPTCEEPGRSLNFDLSVRLADARTVLGGASAH 120
Db 61 MPSAGTLFWVGLICNVNNTCPQPTCEEPGRSLNFDLSVRLADARTVLGGASAH 120
Qy 121 TLAGLGLIATRAAR-----STAQPQPTKQSPLEPPM--- 153
Db 121 SIKDMHKLRLMLRQIKHNSNLKLDQFLVDNFTSGFLQHNLSLPRSTVDSLQANVGLQ 180
Qy 154 -----LDVAELL----- 160
Db 181 KVFLLQGYQLHLASLNGSKLEFIILQGDAEVSALCGLPRKLDAAERVLRYNMDIKPVV 240

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette
DE transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux
DE regulatory protein).
GN Name=ABCA1; Synonyms=ABCL1, CERP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=20345099; PubMed=10884428;
RA Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,
RA Freeman L.A., Cheng J.-F., Osorio J., Renaley A.T., Yang X.-P.,
RA Haendrich C.C., Prades C., Chimini G., Blackmon E.E.,
RA Francois T.L., Duverger N., Rubin E.M., Rosier M., Deneffe P.,
RA Fredrickson D.S., Brewer H.B. Jr.;
RT "Complete genomic sequence of the human ABCA1 gene: analysis of the
RT human and mouse ATP-binding cassette A promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Schwartz K., Lawn R.M., Wade D.P.;
RT "ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
RT regulated by LXR.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RA Qiu Y., Cavellier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies
RT revealing novel regulatory sequences.";
RL Genomics 73:66-76(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Tanaka A.R., Abe-Dohmae S., Arakawa R., Sadanami K., Kidera A.,
RA Kioka N., Amachi T., Yokoyama S., Ueda K.;
RT "A new topological model of functional human ABCA1-signal peptide
RT cleavage and glycosylation of a large extracellular domain.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99194549; PubMed=10092505; DOI=10.1006/bbrc.1999.0406;
RA Langmann T., Klucken J., Reil M., Liebisch G., Luciani M.F.,
RA Chimini G., Kaminski W.E., Schmitz G.;
RT "Molecular cloning of the human ATP-binding cassette transporter 1
RT (hABCL1): evidence for sterol-dependent regulation in macrophages.";
RL Biochem. Biophys. Res. Commun. 257:29-33(1999).
RN [6]
RP SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99364413; PubMed=10431238;
RA Rust S., Rosier M., Funke H., Real J., Amoura Z., Piette J.-C.,
RA Deluc J.-F., Brewer H.B., Duverger N., Deneffe P., Assmann G.;
RT "Tangier disease is caused by mutations in the gene encoding ATP-
RT binding cassette transporter 1.";
RL Nat. Genet. 22:352-355(1999).
RN [7]
RP PHOSPHORYLATION SITES SER-1042 AND SER-2054.
RX MEDLINE=22289331; PubMed=12196520; DOI=10.1074/jbc.M204923200;
RA See R.H., Caday-Malcolm R.A., Singaraja R.R., Zhou S., Silverston A.,
RA Huber M.T., Moran J., James E.R., Janoo R., Savill J.M., Rigot V.,
RA Zhang L.H., Wang M., Chimini G., Wellington C.L., Tafuri S.R.,
RA Hayden M.R.;
RT "Protein Kinase A site-specific phosphorylation regulates ATP-binding
RT cassette A1 (ABCA1)-mediated phospholipid efflux.";
RL J. Biol. Chem. 277:41835-41842(2002).
RN [8]
RP REPRESSION BY ZNF202.
RX MEDLINE=21192304; PubMed=11279031; DOI=10.1074/jbc.M100218200;
RA Porach-Oezuerne M., Langmann T., Heimerl S., Borsukova H.,
RA Kaminski W.E., Drobnik W., Honer C., Schumacher C., Schmitz G.;

RT "The zinc finger protein 202 (ZNF202) is a transcriptional repressor
RT of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene
RL expression and a modulator of cellular lipid efflux.";
RN J. Biol. Chem. 276:12427-12433(2001).
RN [9]
RP INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not
RT ABCG1 via an LXR-independent pathway.";
RL J. Lipid Res. 43:952-959(2002).
RN [10]
RP REVIEW ON VARIANTS.
RX MEDLINE=22790907; PubMed=12763760;
DOI=10.1161/01.ATV.0000078520.89539.77;
RA Singaraja R.R., Brunham L.R., Visscher H., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Efflux and atherosclerosis: the clinical and biochemical impact of
RT variations in the ABCA1 gene.";
RL Arterioscler. Thromb. Vasc. Biol. 23:1322-1332(2003).
RN [11]
RP VARIANTS HDLD2 THR-1091 AND 1893-GLU-ASP-1894 DEL.
RX MEDLINE=20001430; PubMed=10533963;
RA Marciel M., Brooks-Wilson A., Clee S.M., Roomp K., Zhang L.-H., Yu L.,
RA Collins J.A., van Dam M., Molhuizen H.O.F., Loubser O.,
RA Ouellette B.F.F., Senses C.W., Fichter K., Mott S., Denis M.,
RA Boucher B., Pimstone S., Genest J. Jr., Kastelein J.J.P., Hayden M.R.;
RT "Mutations in the ABC1 gene in familial HDL deficiency with defective
RT cholesterol efflux.";
RL Lancet 354:1341-1346(1999).
RN [12]
RP VARIANTS HDLD1 ARG-597 AND ARG-1477, AND VARIANT HDLD2 LEU-693 DEL.
RX MEDLINE=99364411; PubMed=10431236;
RA Brooks-Wilson A., Marciel M., Clee S.M., Zhang L.-H., Roomp K.,
RA van Dam M., Yu L., Brewer C., Collins J.A., Molhuizen H.O.F.,
RA Loubser O., Ouellette B.F.F., Fichter K., Ashbourne-Excoffon K.J.D.,
RA Senses C.W., Scherer S., Mott S., Denis M., Martindale D.,
RA Frohlich J., Morgan K., Koop B., Pimstone S., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Mutations in ABC1 in Tangier disease and familial high-density
RT lipoprotein deficiency.";
RL Nat. Genet. 22:336-345(1999).
RN [13]
RP VARIANTS HDLD1 SER-590; SER-935 AND VAL-937, AND VARIANTS ALA-399 AND
RP MET-883.
RX MEDLINE=99364412; PubMed=10431237;
RA Bodzioch M., Orso E., Klucken J., Langmann T., Boettcher A.,
RA Diederich W., Drobnik W., Barlage S., Buechler C.,
RA Porach-Oezuerne M., Kaminski W.E., Hahmann H.W., Oette K.,
RA Rothe G., Aslanidis C., Lackner K.J., Schmitz G.;
RT "The gene encoding ATP-binding cassette transporter 1 is mutated in
RT Tangier disease.";
RL Nat. Genet. 22:347-351(1999).
RN [14]
RP VARIANTS HDLD1 ILE-929; ARG-597 AND ARG-1477, AND VARIANTS HDLD2
RP LEU-693 DEL; THR-1091; 1893-GLU-ASP-1894 DEL AND LEU-2150.
RX MEDLINE=20540002; PubMed=11086027;
RA Clee S.M., Kastelein J.J.P., van Dam M., Marciel M., Roomp K.,
RA Zwarts K.Y., Collins J.A., Roelants R., Tamasaawa N., Stult T.,
RA Suda T., Ceska R., Boucher B., Rondeau C., Desouch C.,
RA Brooks-Wilson A., Molhuizen H.O.F., Frohlich J., Genest J. Jr.,
RA Hayden M.R.;
RT "Age and residual cholesterol efflux affect HDL cholesterol levels and
RT coronary artery disease in ABCA1 heterozygotes.";
RL J. Clin. Invest. 106:1263-1270(2000).
RN [15]
RP VARIANTS HDLD1 ASN-1289 AND HIS-1800.
RX MEDLINE=20171564; PubMed=10706591;
RA Brousseau M.E., Schaefer E.J., Dupuis J., Eustace B.,
RA Van Berckwegh P., Goldkamp A.L., Thurston L.M., FitzGerald M.G.,
RA Yasek-McKenna D., O'Neill G., Eberhart G.P., Weiffenbach B.,
RA Ordovas J.M., Freeman M.W., Brown R.H. Jr., Gu J.Z.;
RT "Novel mutations in the gene encoding ATP-binding cassette 1 in four

```

RT tangier disease kindreds";
RL J. Lipid Res. 41:433-441(2000).
RN [16]
RP VARIANT HDL1 ASP-1046, VARIANT HDL2 CYS-230, AND VARIANTS LYS-219;
RP ILE-825; MET-883 AND LYS-1587.
RX MEDLINE=20396633; PubMed=10938021;
RA Wang J., Burnett J.R., Near S., Young K., Zimman B., Hanley A.J.G.,
RA Connelly P.W., Harris S.B., Hegelle R.A.;
RA "Common and rare ABCA1 variants affecting plasma HDL cholesterol.";
RA Atheroscler. Thromb. Vasc. Biol. 20:1983-1989(2000).
RL [17]
RN VARIANT HDL1 TRP-587, AND VARIANT LEU-2168.
RP MEDLINE=21157002; PubMed=11257260;
RX Bertolini S., Pisciotto L., Seri M., Cusano R., Cantafora A.,
RA Calabresi L., Franceschini G., Ravazzolo R., Calandra S.;
RA "A point mutation in ABC1 gene in a patient with severe premature
RT coronary heart disease and mild clinical phenotype of Tangier
RT disease.";
RL Atherosclerosis 154:599-605(2001).
RN [18]
RP VARIANTS LYS-219; MET-883 AND ASP-1172.
RX MEDLINE=21157003; PubMed=11257261;
RA Brousseau M.E., Bodzioch M., Schaefer E.J., Goldkamp A.L., Kielar D.,
RA Probst M., Ordovas J.M., Aslanidis C., Lackner K.J.,
RA Bloomfield Rubins H., Collins D., Robins S.J., Wilson P.W.F.,
RA Schmitz G.;
RA "Common variants in the gene encoding ATP-binding cassette transporter
RT 1 in men with low HDL cholesterol levels and coronary heart disease.";
RL Atherosclerosis 154:607-611(2001).
RN [19]
RP VARIANT HDL1 LEU-1506.
RX MEDLINE=21369433; PubMed=11476961; DOI=10.1016/S0925-4439(01)00053-9;
RA Lapicka-Bodzioch K., Bodzioch M., Kruehl M., Kielar D., Probst M.,
RA Kiec B., Andriukovics H., Boettcher A., Hubacek J., Aslanidis C.,
RA Suttorp N., Schmitz G.;
RT "Homogeneous assay based on 52 primer sets to scan for mutations of
RT the ABCA1 gene and its application in genetic analysis of a new
RT patient with familial high-density lipoprotein deficiency syndrome.";
RL Biochim. Biophys. Acta 1537:42-48(2001).
RN [20]
RP VARIANTS HDL1 ASN-1289 AND TRP-2081, AND VARIANT LYS-219.
RX MEDLINE=21369433; PubMed=11476965; DOI=10.1016/S0925-4439(01)00058-8;
RA Huang W., Moriyama K., Koga T., Hua H., Agata M., Kawabata S.,
RA Mawatari K., Imamura T., Eto T., Kawamura M., Teramoto T., Sasaki J.;
RA "Novel mutations in ABCA1 gene in Japanese patients with Tangier
RT disease and familial high density lipoprotein deficiency with coronary
RT heart disease.";
RL Biochim. Biophys. Acta 1537:71-78(2001).
RN [21]
RP VARIANTS LYS-219; ALA-399; MET-771; PRO-774; ASN-776; ILE-825;
RP MET-883; ASP-1172; LYS-1587 AND CYS-1731.
Query Match 51.6%; Score 5751.5; DB 1; Length 2261;
Best Local Similarity 50.2%; Pred. No. 0;
Matches 1145; Conservative 367; Mismatches 609; Indels 161; Gaps 26;
Qy 1 MAFVTQMLLWKNFMVRRPQVOLLVLLWPLFLPFLVAVRSHHPLEHCHFPNKP 60
Db 1 MACPQRLLLWKNLTERRQTCQLLEAVAMPFLIFLILISVRLSYPPYECHFPNKA 60
Qy 61 LPSAGTVPWLOGLICNVNNTCFPOLTPCEBGRSLNENDSLVSLLDARTVLGASAH 120
Db 61 MPSAGTLPWOGIICNANNPCFRYPTEAGPVGVGNFNKSVARLFSDDARLLYSOKDT 120
Qy 121 TLAGLGLKIATLRAARSTAQ-----POPTKOSPLEPPML-- 154
Db 121 SMKDMRKVLRTLOQIKKSSNKLQDFLVNDNETSGFLYHNLSPKSTVDKMLRADVILH 180
Qy 155 -----DVDEL-----LTSLL 164
Db 181 KVFLLQGYQLHLTSLCNGSKSEMIQLGQDVSELCGLPREKLAARVLRNSMDILKPI 240
Qy 165 RT--ESIGLALGOAQEPHLSLEAEADLAQELLALRSIVEL---ALLQPRGTSGPLE 219

```


Db 1307 LLSGMDGKSGYQVKGWKLTKQOQFVALLWKRLLIARRSRKGFPAQIVLPAVFVCIALVFSL 1366
QY 1259 IVPFPGHYPALRLSPMTYGAQVSFTSEADAPGDPGRARLEALLQBAQ-----LEE 1308
Db 1367 IVPFPGKPSLEQPMWNEQYTFVSNDAPEDTGTELLNLAUTKDPGFGTRCKEGNPIPD 1426
QY 1309 PVPQSHSRFSAPAEVAKVLAGSNMTTPESPSPACQCSQFGARRLLPDCPAAAGGPPP 1368
Db 1427 TPCQAGEEBWTAPVQPTIMDLFGNWTMQNPSPACQCSQDKIKKMLPVCPPGAGGLPP 1486
QY 1369 POAVTSGSEVQNLGRNLSDLVKTYPRLVQGLKTKWNEVRYGFSIG-GRDPGLP 1427
Db 1487 PORKQNTADILQDLGRNLSDLVKTYQIIRAKSLNKITWNEPRYGFSGVSNQALP 1546
QY 1428 SQEGLGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAOSLKIWFNNKGWHSWAPV 1487
Db 1547 PSEVNDATKQMKHKLAKDSSADRFNLNLSGRFMTGLDTRNNKVFNNKGWHAISFL 1606
QY 1488 NRASNAILRAHLPPOGPARHAHSITTLNHLPLNLTKEQLGEAALMASSVDVLVSIQVVFAMS 1547
Db 1607 NVNNAILRANLQKGENPSHYGITAFNHLPLNLTQQLSEVAPMTTSDVLVSIQVIFAMS 1666
QY 1548 FVPASFTVLIERVTRAKHLQMLGSLPTLYWLNFLWDMCNVLPACIVVLIFLAFQO 1607
Db 1667 FVPASFTVLIERVTRAKHLQMLGSLPTLYWLNFLWDMCNVLPACIVVLIFLAFQO 1726
QY 1608 RAYVAPAMLPALLLLLLYGNWITPLMPASFFSPSTAYVVLTCINLFIGINGSMATF 1667
Db 1727 KSYVSTNLPLVALLLLYGNWITPLMPASFFSPSTAYVVLTCINLFIGINGSMATF 1786
QY 1668 VLELFSQDLQVRSRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPLRW 1727
Db 1787 VLELFTDNKLNINDILKSVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPLRW 1846
QY 1728 EVWGKILLAMVQGLPFLFTLLQHRSQLLPQVRSLPLGDEDEDAVERERVOGA 1787
Db 1847 DUVGRNLFAMAVEGVVFLITVLIQYRFFIRPPVNAKLSPLNDEDEDAVERERVOGA 1906
QY 1788 TQGDVLVRLNLTQVYRGORMPAVDRLCLGIPGECFGLLVGNAGKSTFRMTVTDTLAS 1847
Db 1907 GQNDILEIKELTKIYRRKRPVADRI CVGIPGECFGLLVGNAGKSTFRMTVTDTLAS 1966
QY 1848 RGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVOTAG 1907
Db 1967 RGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVOTAG 2026
QY 1908 SGLARLGLSWADRPAGTYSNGKRLATALVGDPAVFLDEPTTGMDSARRFLWNS 1967
Db 2027 WAIRKGLGLVKGEGYAGNYSNGKRLATAMALIGGPPVFLDEPTTGMDSARRFLWNS 2086
QY 1968 LLAVREGRSVMLTSHSMECEALCSRLAIMVNGFRCLGSPHLLKGRFAAGHTLTLRVP 2027
Db 2087 ALSVVKEGRSVMLTSHSMECEALCSRLAIMVNGFRCLGSPHLLKGRFAAGHTLTLRVP 2146
QY 2028 AARS--QPAAAFVAAEFCELSREAHGRLRQLPFGGRCALARYVGLAVHGAEBHVED 2085
Db 2147 GSNPDLKPQDFGLAFPGSPKERNMLQYL--PSSLSLARLFSILSQSKRELHIED 2205
QY 2086 FSVSOTMLEVFLYFSKQKDE---DTEBQKAGVGDVDPAGLOHPKRVQOFLDDPSTA 2142
Db 2206 YSVSOTMLEVFLYFSKQKDE---DTEBQKAGVGDVDPAGLOHPKRVQOFLDDPSTA 2257
QY 2143 ET 2144
Db 2258 ES 2259

RESULT 10

Q80ZB2

ID Q80ZB2 PRELIMINARY; PRT: 2201 AA.

AC Q80ZB2;

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ATP-binding cassette 1.
GN Name=Abcal;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Ananthanarayanan M., Mirza M.F.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
EMBL; AY208182; AAC53557.1; .
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran.; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2201 AA; 246553 MW; B1472978BFC3E688 CRC64;

Query Match 49.5%; Score 5510.5; DB 2; Length 2201;
Best Local Similarity 49.5%; Pred. No. 0;
Matches 1108; Conservative 361; Mismatches 575; Indels 195; Gaps 28;

QY 61 LPSAGTVPWLOGLICNVNNTCPQOLTGCEEGRLSNFNDSILVSRLLADARTVLGASAH 120
Db 1 MPSAGTLPWQGIICNANNPCFRYPPTGCAFGVGNFNKSIVSRFLSDAQRIILYSQKT 60
QY 121 TLALGLKLIATLRAAR-----STAOPQTKQSPLE----- 150
Db 61 SIRDHKKVLRTLQIKHPNSNLKLDQFLVDNETSGFLQHSLSLPSRVAVDNLLQADVSLQ 120
QY 151 -----PMLDVAELL----- 160
Db 121 KVFLQGYQLHLASLNGSKLEIIRPDLKVSALCSLPREKLDAPELRSNMDILKPM 180
QY 161 -----TSLRPTESLGLALQAOEPLHSLLEAEADLAQELLALRSVLRALLQ-----RP 210
Db 181 TKLANSTSLLPQHILA-----EATTTLLDSLGLAQELFSTKSWDMRQEVMTN 233
QY 211 RGTSGPL-ELLSEALCSVRGSPSTVGPSLNWEASDLMELVG---QSPESALPDSSLSL 265
Db 234 SGSTQIYQAVSRIVCGHPEGGLKIKSLNWEYDNNYKALFGNGTEDEDTTFYDNSTTP 293
QY 266 ACSBELGALDSHPILSRLLWRRLKPLILGLKLPADPTPTFRKLMAQVNRTPBELTLRLDVR 325
Db 294 YCNLMKNLESSPLSRIIWKALKPELLIKILYTPDTPATQVMAEVNKTFOELALFPDLE 353
QY 326 EVMELGPRIETFWNDSNVAMLQRLQMQ--DEGRQPRPGGRDHMEALSFL----- 377
Db 354 GMBEELSPQIWTFFMESSEQMDLVRPMLDLRQNDQFWERKLDGLYWTAQDMAFLAKNPED 413
QY 378 --DPGSGGYQWQDAHADVGLVGLGRVTECLSLDKLEAAPSEALVQLQLAEHFRFW 435
Db 414 VQSPNGSVYTWREAFNETNQAIQIISRPMECVNKLKLEPIFTEVTLINKSMOLLDAKFS 473
QY 436 AGVFLGPEDSSDPTHTPDLG--PGHVRIKIRMDIDVVTTRNKIRDRFPWDPGPAADPL 493
Db 474 AGIDFTG-----ITPDSVELPHHVKDKIRMDIDNVERTNKIKDGYWDPGPAADPF 523
QY 494 TDLRYVWGGFYVLQDLVERAAVRVLSGANPRAGILYLOQMPYPCYVDDVFLRVLRSPLDF 553
Db 524 EDMRYVWGGFYVLQDLVERAAVRVLSGANPRAGILYLOQMPYPCYVDDVFLRVLRSPLDF 583

Db 531 EIQLTQRAISLLEENRFWAGVFP-----PDMYPWTSALPTHVKYKIRMDIDVVEKTN 582
Qy 478 KIRDRFWDGPAADPLTDLRYVWGGFVYLDQVRAAVRVLSGANPRAGLYLOMPYPCY 537
Db 583 KIKDRYWDGPRADVEDFEDRWGFAFYLDQMIHQGITSQAQVXKVPVGIYLOMPYPCF 642
Qy 538 VDDVFLAVLSBSLPLFLTILAWIYVTLTKAVVREKETRLDTRWANGLSRAVLWGLF 597
Db 643 VDSFMIILNRCFFIFMWLAWIYVSMTVKSIVLEKELRLKETLKNQGVSNVITWCTWFL 702
Qy 598 SCLGPFLLSALLVLVLLKGLDILYSHVGVVFLFAAFVAVATVQSFLLGAFFSRANLAA 657
Db 703 DSPSIMSNSIFLLTIFIMHGRIILYHNSPFLIFLLAFSTATINQCFLLSTFFSRASLAA 762
Qy 658 ACGLAYFSLYPLVYLCVANDRLPAGRVAASLLSPVAFGFCESLALLEEQEGAGWH 717
Db 763 ACSGVYFTLYLPHILCFAMQDRMTADLKNVAVLSLSPVAFGFGTEYLARFEEQGLQWS 822
Qy 718 NVGTRP-TADVFLSAQVSGLLILDAALYGLATWYLEAVCGOYCIPEPWNFPFRSRWCG 776
Db 823 NIGKSPMEGDEFSFMSKMMMLDAILYGLLAWYLDQVFPNGYGTPLFWYFLQESYWL 882
Qy 777 -----PRPKSPAPCPTPL-DPK-----VLVEEAPGLSPGVSRSLKRPFGSP 820
Db 883 GEGCSTREERALEKTEBITEMEDPEHPEGINDAFFERELPGLVPGVCVNLVIFEPYS 942
Qy 821 OPALUGLSLDPYQGHITAFIAGHNAGAKTTLSILSGLFPSPGGSFAFIGHDVSRSSMAAIR 880
Db 943 RPAVRLNITFYENQITAFIAGHNAGAKTTLSILSGLFPSPGGSFAFIGHDVSRSSMAAIR 1002
Qy 881 PHLGVCQYVNLFDMLTWDHWHVYGRILKGLSAAVGPEDQRLLODGLVSKQSVQTRHL 940
Db 1003 QSLGMCPOYNILFPHLTVAEHILFYAQLKGSWEAQLEAMEAMLEDTGLHKKRNEEAQDL 1062
Qy 941 SGGMOKLSVAIAFVGGSVQVILDEPTAGVDPASRRGIWELLKYRGRTLLISTHLE 1000
Db 1063 SGGMOKLSVAIAFVGGAKVILDEPTSGVDYPSRSIWDLLLYKRGRIIIMSTHWE 1122
Qy 1001 AELGDRVAVVAGRLCCGSPFLRRHLGSGYVLTILVKARLPLTNEKADTMEGSDVT 1060
Db 1123 ADLLGDIRAIIISQGRILCSGTPLFKNCFGTGFYLTIVR---KMKNIQSQTCEGTCSC 1179
Qy 1061 ROEKKNSQSGRVG--TP-----QLLALVOHVVPGARLVEELPHELVLVLPYTGAAH 1109
Db 1180 ASEGFSTRCPAHIDEITPEQVLDGVDNMDMVQHVPEAKLVEICIGQELIFLLPNKFK 1239
Qy 1110 DGSFATLFLRELDRLAELRTGYISDTSLEEIEFLKVVECAADTMDGSCGQHLCTGI 1169
Db 1240 QRAYASLFRELEDLADLGLSSFGISDTPLEIEFLKV-----EDSDSGPLFAGGT 1290
Qy 1170 AGLDVTURLK-MPQETALE-----NG---EPAGSAPETQGGSPDAVRVQ-GWALTRQ 1219
Db 1291 QOKRENLRHPWLSPREKARQIPQGSNGCSKEPA-PHEGQPSSEPEARSRLNTGAOLIIQ 1349
Qy 1220 QLOALLKRELLARSRGFLPAQIVLPAFVGLALVFLSLVPPCHYPALRLSLTMYGQAO 1279
Db 1350 HVQALLVKRFHTTRSHKOFIAQIVLPATFVFLALMLSIIIVPPGEYPALILHPWYGOQ 1409
Qy 1280 VSPFSEDAPGDPGRARLEALLQOAG-----LEEPPVQHSRHSFAPEVPAEPAVAK 1329
Db 1410 YTFESLQDQSEQLAALADVLNLPFGNGRCLKEWLPPEYPCGNST-PWKTSPSVNITH 1468
Qy 1330 VLASGNWTPSPSPACQSQGARLLPDCPAAAGGPPPPQAVTSGSEVQNLTRNLSD 1389
Db 1469 LFQKQOWTPEKPSRCSTREKLTMLPECEGAGGLPPQRIQRSTEILQDLTNRNIS 1528
Qy 1390 FLVKTYRLVROGLKTKKWNVEYGGFSLGCRDPGLP-SQOELGRSVEELWALLSLPG 1448
Db 1529 FLVKTYPALTRSSLSKSPWNEQYGGISIGGKLPILPITGEALVEFLSHLQIMN-VSG 1587
Qy 1449 GALDR-VLKNLTAWAHSIDAQSLKIWFNNKGMHSMVAFVNRASNAIILRAHLPPGPARHA 1507

Db 1588 GPITREASKEMPAFLKHELETEDNIKVFNKNGHVALVSFLNVAHNALTRSLHKDKNPEE 1647
Qy 1508 HSTITLNLPLNLTKEQLSEAAALWASSVDVLVSVICVFPAMSFVPASFTLVILIERVTPAKH 1567
Db 1648 YGITVISQPLNLTKEQLSEITVLTASVDAVAICVIFAMSFVPASFTLVILIERVTPAKH 1707
Qy 1568 LQJMGGLSPTLYWLGHNFLNFCNVLVPACIVLILFLAFOQRAVVPANLALLLLLYG 1627
Db 1708 LQFVSGVSPTYWLTWFLDMIMYASAAUVGIFGVGKQKAYTSPENLPAIALLMLYG 1767
Qy 1628 WSTPLMYPASFFSPSTAYVVLTCINLFIGINGSMAFVLEFSDQ-KLOEVSRLKQ 1686
Db 1768 WAVIPMWYPASFLFDVPSTAYVALSCANLFIGINSSAITFELFENNWTLLRFNAMLK 1827
Qy 1687 VFLIPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWEVVGKLLAMVQGPFL 1746
Db 1828 LLIIIFHFCLGRGLIDALSQAQVTVYARFGEESHSTNPFQMDLIGKLVAMAAEGVVYLL 1887
Qy 1747 FTLLLOHR---SOLLPOPRVRSPLLGEEREDVARERERVVOGATQGDVLVLRNLTKVYR 1803
Db 1888 LTLFIQHHPFLTRWSEPAKE--PII-DEDDVAERQRIISGKNKTDILRLNELTKIYS 1944
Qy 1804 GQMPAVDRCLGIPGCECFGLLVNGVAGKSTFRMVGTDLASRGSAVLGHSVAREPS 1863
Db 1945 GTSSPAVDRLCVVRGCECFGLLVNGVAGKSTFRMVGTDLASRGSAVLGHSVAREPS 2004
Qy 1864 AAILSWGYCPOSDAIFELLTGREHLELLARLGRVPEAQVATAGSLARLGLSWYADRP 1923
Db 2005 DVHQSMGYCPOFQVAVDVLTLGREHLYLARLGRVPADEIERVAVNSIQSLGLSLYADRLV 2064
Qy 1924 GTYSGGNKRLATALVGPVAVVFLDEPTTGDPGARREFLNLSLLAVVREGSRVMTSH 1983
Db 2065 GTYSGGNKRLATALVGPVAVVFLDEPTTGDPGARREFLNLSLLAVVREGSRVMTSH 2124
Qy 1984 SMEECALCRLAIMVNGFRCLGSPQHLKGRFAAGHTLTLRVPAAARS-----QPAAAF 2037
Db 2125 SMEECALCRLAIMVNGFRCLGSPQHLKGRFAAGHTLTLRVPAAARS-----QPAAAF 2184
Qy 2038 VAEPFGSEUREAHGRLRQLPFGRCALARVFGELAVHGAHGVEDFVSQTMLEEVF 2097
Db 2185 FQGNFPGSVQRRHYNMQFQV---SSSLARIFRLLISHKSLLEEYSVTQTLDOVF 2241
Qy 2098 LYFSKQDQKDED 2109
Db 2242 VNFAKQOTETHD 2253

RESULT 14
AAR87836 PRELIMINARY; PRT; 2269 AA.
AC AAR87836; 27, Created
DT 20-MAY-2004 (T-EMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (T-EMBLrel. 27, Last sequence update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.;
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.";
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY427779; AAR87836.1; -

SQ SEQUENCE 2269 AA; 256491 MW; 9B6E349FF17895A1 CRC64;

Query Match

Best Local Similarity 45.1%; Score 5029; DB 2; Length 2269;

Matches 1046; Conservative 358; Mismatches 666; Indels 222; Gaps 37;

QY 1 MAFVQLMLLWKNFYRRQPVQLLVLLWPLFFILVAVRHSHPPLEHCHFPNKP 60
DB 1 MGFARQIQLLWKNWTLKQKIRFVVELWPLSLFLILILWLNINPLYSQHECHFPNKA 60
QY 61 LPSAGTVPLWGLICNNVNTCPQITPGPEGRGLNFDNLSVRLADARTVVGASAH 120
DB 61 MFSAGMLPWLQMFNCVNNPFCQNPPTGSGPGIVSNYNNSILARVFRDFOELLDAPE 120
QY 121 TLAGLGKLTATLRAARSTAPQPTQKSPLEPPMLDAE----- 158
DB 121 HFGHVWKEFQTLRLMDTLRTHPVRVAGRGIRVDLKDDEETLTLFLMKNIGLSDSVYL 180
QY 159 LTLTLRLTESL-----GLALGOAO-- 177
DB 181 LLSQVRPEQFAHGPDLMLKDIACTETLLERFIIFSQRGAQTVRDAMCSLSQTLQWV 240
QY 178 -----EPHLSLEAEADLAQELALRS-----LVELRALQRP----- 210
DB 241 EDTLYANVDFFKLFRVLPFLDSSSQINLRSGRVSFDSISSRIREFIHRPSVEDLLWVT 300
QY 211 ---RGTSGP-----LELLSEALCSV-RGFSSTVGPSLNMYEASDLMELVGQEPESALP- 259
DB 301 KPLTQTGGPETAQMLISLSDLCGYPEGGGRVF-SFNWYEDNNYKAFGLGIDSTRKDOI 359
QY 260 ---DSSLSPACSGALGDHSLRLLWRLKPLILGKLLFAPDTPFTRKMAQVNRTPF 316
DB 360 YSYDKRTTFNCALLQSLESNPLTKIWAAPLWGLKILFTDPAVRRIILONANSTFE 419
QY 317 EUTLRDVRVWEMIGPRIFTFMDPSSNNVAMLORLL-----QMDEGRQPRPG 365
DB 420 ELERLKLKWAEEVGPQIWFDFRSTQMTMRDLENTPVKGLFNSQLGEGIT----- 474
QY 366 GRDHMEALRSFLDPG-----SGYSWQDAHADVGHVGLTGRVTECLSDLDKLEAAPS 417
DB 475 -----AEAMLNFLHKGPRSQADDMANFQWRDVFNTDRTLRTSKYLECLILDKEFSYDD 530
QY 418 EAALYSRALQLLAEHRFAGVYFLGPDSSDTEHPTDGLGPHVRIKIRMDIDVYVTRIN 477
DB 531 ELQLTQALSLLEENRFAGVVF-----PDMYPWTSALPTHVKYKIRMDIDVVEKTN 582
QY 478 KIRDPWDGPAADPLTLDRYVWGFVYQLDIVERAAVRVLSGANPRAGLYLQMPYPCY 537
DB 583 KIKDRYWDGPRADPVEDPRYTWGGFAYLQDMIEQGITRSQAQVXVPVGVYLYQMPYPCF 642
QY 538 VDDVFLVLSRLPLFLTLAWIYSVTLTKAVVREKETRLRMTAMGLSRAVLWGLWFL 597
DB 643 VDDSEMIILNRCFFIPWLAWIYSVMTKSVLEKELRLKETLKNQGVSNVAVICTWFL 702
QY 598 SCGLGPFLLSAALLVLVLGDLILYSHPGVWFLFAAFVAVTVTQSFLLSAPFSRANLAA 657
DB 703 DSFSIMSWSIPLTTFIMHGRILHVSNNFIFLFLAPSTATIMQCFLSTFFSASLAA 762
QY 658 ACGGLAYFSLYPLVYLCVAVMRDLRPPAGGRVAASLSPVAFGCGCSLALLEBQSGAOWH 717
DB 763 ACSGVITYFTLYLPHILCFPAWQDRMTADLKNVLSLLSPVAFGFGTBYLARFEBEQGLQWS 822
QY 718 NVGTTP-TADVPSLAQVSGLLILLDAALYGLATWYLEAVCPGQYGTPEWNPFFRSYWG 776
DB 823 NIGKSPMEGDEFSFLMSMKOMMLDAALYGLLAWYLDQVPPGNYGTPPLPWYFLLOSXYLG 882
QY 777 -----PRPPKSPAPCTPL-DPK-----VLVEBAPPLSGVSVLSLEKRFPGSP 820
DB 883 GEGCSTRERALEKTEPTTEEMEDPEHPGINDAFERELPGLVPGVCVKNVLKLFEPYS 942
QY 821 QPALRGLSLDFYQGHITAFLGHNGAGKTTTLLSGLFFPPSGGSFAFILGHDRSSMAAIR 880

DB 943 RPAVDRLNITFYENQITAFLGHNGAGKTTTLLSILGTPPTSGTVLIGKDIETSLDAVR 1002
QY 881 PHLGVCPOYNVLPDMLTYDEHWFYGRLLKGLSAAVVGPEQDRLLODVLGVSQSVQTRHL 940
DB 1003 OSLGCMCPQTNILFHHLTVAEHLFVAQLKLGKSWEBAQLEMEALBETDGLHHRKNEBAQDL 1062
QY 941 SSGMQKLSVAIAFVGGSQVWILDEPTAGVDPASRRGIWELLKLYKREGRTLLSLTHHDE 1000
DB 1063 SSGMQKLSVAIAFVGGAKVILDEPTSGVDPYSRRSIWDLKLLKYSRGTIIMSTHMD 1122
QY 1001 AELLGDRVAVAGGRGCCGSPFLFURRHGSGYYITLVKARLPLTTNEKADTDMGSDVT 1060
DB 1123 ADLLGDRIAIISQGRSLYSGTFLKNCFGTGYTLVR---KMKNIQSORTCEGTSC 1179
QY 1061 ROEKNGSGQSRVG--TP-----QLLALVQHVMPGARLVELPHELVLVLPYTGAAH 1109
DB 1180 ASEGSTPCPAHIDITPEQVLDGVDNBLMDMVQHVHPEAKLVEGIGELIFLFLKNKPK 1239
QY 1110 DGSFATLRELDTRLAEIRLTGYGIDSTLSBEIFLKVVVECAADTDMEDGSCGQHLCTGI 1169
DB 1240 QRAYASLFRELEDTLADLGLSGFISDTPLEIFLKV-----EDSDSGPLFAGGT 1290
QY 1170 AGLDVTBLK-MPPOETALE-----NG---BPAGSAPETDQSGDPDVGVRVQ-GWALTRQ 1219
DB 1291 QOKRENLRHPWLSPREKARQIPQSGNSGCSXBP-PPHEGQPSSEPEARSLMTGAQLITQ 1349
QY 1220 QLOALLKRLFULLARRSREGLFAQIVLPALFVGLALVFSIIVPPGHYPALRLSPMYGAQ 1279
DB 1350 HVQALLVXRFHHTIRSHKDFLQIAIVLPATVFLPALMLSIIVPPFGYFALILHPMYGQ 1409
QY 1280 VSFFSEADPGDPPGRARILLEALLQEAQ-----LSEPPVQHSRHSFSAPEVPAEVAK 1329
DB 1410 YTFESLDQPSQEQLAALADVLNKGFCNRCLEKMLPEYPCNST-PWKTPSVSPNITH 1468
QY 1330 VLASGNWTPSPSPACQSPGARLLPDCPAAAGPPPPQAVTSGSEVVQMLTGRNLS 1389
DB 1469 LFQKQWTPKPSPCSCRCSTREKLTMLPECEGAGLPQRIQSTELQDLTNRNID 1528
QY 1390 FLVKTYPPLVRQGLTKKMWNEVRYGGFSLGGRDPLP-SQOELGRSEVELWALLSLPLG 1448
DB 1529 FLVKTYPALIRSSLSKSKFWNEQRYGGISIGGKLPILITGEALVEFLSHGQIMN-VSG 1587
QY 1449 GADLR-VLKNITAWAHSILDAQDSLKIWFNNKGMWMAFVNRASNAIRLAHLPPGPARHA 1507
DB 1588 GPIFREASKEMPAFLKLETDENIKVWNNKGWALVSFLVNAHNAIRLTSLHKDKNPE 1647
QY 1508 HSITTLNHLNLTKEQLSEALMASSVDVLSICVVPAMSFVPASFTLVLIEERTVRAKH 1567
DB 1648 YGITTVISQPLNLTKEQLSEITVLTAADVAVAIQVIFAMSFVPASFVLYLIQERVNKAKH 1707
QY 1568 LQLMGSLSPTLVWLGNFMDMNVLPACIVVLIFLAFQORAYVAPNIPALLLALLLVG 1627
DB 1708 LQFVSGSPPTTYLWTLNLDIMNVAVSAALVVGIVPGFKKAYTSPENLPALIALMLVG 1767
QY 1628 WSITPLMYPASFFSVSPSTAYVLTCLINLTGINGSMAITVLEIFSDQ-KLQEVSRILKQ 1686
DB 1768 WAVIPMYYPASFLDVPSTAYVALSCANLFIGINSSAITFILEFNWNTLLRFRNALRK 1827
QY 1687 VFLIPPHCLGRGLIDMVNRQAMADAFERLGDROFQSPPLRWEVVGKNLWAMYQGLFL 1746
DB 1828 LLIIPPHCLGRGLIDLALSQAQVDVYARFGEHSTNPFQWDLIGKLVMAAAEGVWVLL 1887
QY 1747 FTLLLOHR---SOLLPOPRVRSPLGLGDEEDVARERERVQATGQGVILRLNLTKYR 1803
DB 1888 LTLFIQHFFHFRWVSEPAKE--PII-DEDDVABEERQRIISGGNKTDLIRLNEUTKIYS 1944
QY 1804 QORPAPVDRCLGIPPGCEFCGLLVNGAGKTTSTFRVMTGDTLASRGEAVLAGHSVAREPS 1863
DB 1945 GTSSPAVDRLCVGVRPGCEFCGLLVNGAGKTTTFKMTGDTTIVTSGDATIAGKSILTNIS 2004
QY 1864 AAHLISMGYCPQSDAIFELLTGREHLELRLARGVPEAQVQTAGSGLARLGLSWYADRP 1923

Db 2005 DVHOSMGYCPQFQDAVDDLLTGREHLVLYARLGRVPADEIERVANWSIQSLGLSLYADRLV 2064
Qy 1924 GTYGGNKRKATLALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGSVMLTSH 1983
Db 2065 GTYGGNKRKSLSTAIALMGCPPLVLLDEPTTGMDPQARRMLWNTIVSIIREGRAVILTSH 2124
Qy 1984 SMESCEALCSRLAIVNGRFRCLSGPOHLKGRFAAGHTLTLRVPAARS-----QPAAAF 2037
Db 2125 SMEECEALCTRLAIVMGKTFQCLGTIOHLKYKFGDGYIVTWKIKSPKEDLLPLNPVEQF 2184
Qy 2038 VAAFFPFSSELRAGHGRRLRFQPPGRCALARVFGELAVHGAEGVEDFSVSQTMLEBEVF 2097
Db 2185 FQGNFPGSVQERHYNMLOFQV---SSSSLARI FRLLSHKDSLLEI BEYSVTQTLDOVF 2241
Qy 2098 LYFSKOQKQKDE 2109
Db 2242 VNAFAKQQTETHD 2253

RESULT 15
Q6T942
ID Q6T942 PRELIMINARY; PRT; 2269 AA.
AC Q6T942;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.;
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.";
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY427777; AAR87834.1; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005951; Rim_ABC_transpt.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR01257; rim protein; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2269 AA; 256482 MW; 00AF870E6D9BE9A5 CRC64;

Query Match 45.1%; Score 5028; DB 2; Length 2269;
Best Local Similarity 45.6%; Pred. No. 1.3e-292;
Matches 1046; Conservative 357; Mismatches 667; Indels 222; Gaps 37;
Qy 1 MAFVQMLLWKNFMYRRQVQLLVLLPLFLFFILVAVRSHHPLEHHECHFPNKP 60
Db 1 MGFARQIQLLWKNWTLRKIRFVVELVPLSLFLILWLNRINPLYSQHECHFPNKA 60
Qy 61 LPSAGTVPWGLICNVNNTCPQPTGCEPGRLSNFNDLSVSLRLADARTVLGASAH 120
Db 61 MPSAGTVPWGLICNVNNTCPQPTGCEPGRLSNFNDLSVSLRLADARTVLGASAH 120
Qy 121 TLGAGKLIATIAARASTAQOPTKQSPLEPMPDLDAE----- 158
Db 121 HFGHWKFEQTLRLSLMDLTLRTHPVRVACGRIRIRDLVLDDEETLTLFLMKNLGSDSVVYL 180

Qy 159 LLTSLRTESL-----GLALGOAQ-- 177
Db 181 LLNSQVRPEQFAHGVDPMLDKIACSETLLERFIIFSQRGAQTVRDAMCSLSGTLQWV 240
Qy 178 -----EPLHSLLEAAEDLAQELLALRS-----LVELRALLQRP-- 210
Db 241 EDTLYANVDFKFLRVLPPTLLDSSSQGINLRSNGRVFSDISSRIRERFIHPSVEDLLWVT 300
Qy 211 ----RGTSGP-----LELSEALCSV-RGPGSVTVGPSLWYEASDLMELVQGPESALP- 259
Db 301 KPLTQTGGTETFAQLMSILDLCCGYPEGGSRVF-SFNWYEDNNYKAFGLIDSTRKDPI 359
Qy 260 ----DSSLSPACSELIGALDHPSLRLLWRRLKLLGLKGLFADPTPTTKLMAQVNTFE 316
Db 360 YSYDKRTTTTCNALIOSLESNPLTKIAWRAAKPLVMGKILFTPDSPAVRILQANSTFE 419
Qy 317 ELTLRLDRVREVMELGPRIETFMNDSNVAMLOKLL-----QMODEGRRQPRPG 365
Db 420 ELERLKLKVAWEVGPQIWIYFDRSTQMTMIRDITLENPTVKGLNSQLGEGIT----- 474
Qy 366 GRDMEALRSFLDPG-----SGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPS 417
Db 475 ----AEAMNLFHKGPRESQADDMANPMDRDNITDRTLRLTSKYLECLILDKFESYDD 530
Qy 418 EAAVSRALOLLAHREHWAGVFLGPDSDPTTEHPTPDGLPGHVRKIRBMDIDVTRTN 477
Db 531 EIQLTQRLSLLEENRFWAGVF-----PDMYMTSALPHTHVKYKIRBMDIDVBEKTN 582
Qy 478 KIRDRFWDGPAAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAGILQOMPYPY 537
Db 583 KIKDRYWDSPRADPVEDFRYINGGFAYLQDMLEQGITRSQAQVEVPGVYLQOMPYPCF 642
Qy 538 VDVFLRVLRSILPLFTLAWISYVTLTKAVVREKETRLDRTNRAMGLRAVLWLQWFL 597
Db 643 VDDSFMIILNRCPPIFMVLAWISYVSTVKSIVLEKELRLKETLKNQGVSNVITWTFWL 702
Qy 598 SCGLGPFLLSAALLVVLKGLDILPYSHPGVVFELFAFAVATVTSQFLLSAFTSRANLAA 657
Db 703 DSFSIMSMSIFLTLTIFIMHGRILHYSNPFILFLFLAFSTATIMQCFLSTFFSRASLAA 762
Qy 658 ACGLAYFSLYLPVLCVAMWRDLRPAAGRVAASLLSPVAFGFCESIALLEEQEGEAQWH 717
Db 763 ACSGVITYLYLPHILCFACQDRMTADLKWAVSLLSPVAFGFGTEYLARFEEQGLQWS 822
Qy 718 NVGTRP-TADVSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPWPMPFRSRWCG 776
Db 823 NIGKSPMEGDEFSPLMSKMWMLDAAALYGLLAWYLDQVFPNGYGTPLPFWYFLLOESYWL 882
Qy 777 -----PRPKSPAPCTPL-DPK-----VLVEEAPGLSPGVSVRSLEKRFPGSP 820
Db 883 GEGCSTREERALEKTEPITEEMEDPEHPGINDAFFERELPGLVPGVCVKNLVKIFPYS 942
Qy 821 QPALRGLSLDFYOGHITAFLGHNGAGKTTTLLSLGLFPSPGGSFAFILGHHDVRSMAAIR 880
Db 943 RPAVDRLNITYFENOITAFLGHNGAGKTTTLLSLTGLLPSTGTVLIGKDIETSLDAVR 1002
Qy 881 PHLGVCPOYNVLFDMLTVDVSHVWFYGRKLGSLAAVVGPEQDRLLQDVLGVSQVQTRHL 940
Db 1003 QSLGMCPOYNILPHLTVAEHILFYAQLKGSWEAAQLEAMEAMLEDGLAHKRNEEAQDL 1062
Qy 941 SGGNORKLSVAIAVPGSOVVIDDEPTAGVDPASRRCIGIWELLKLYRSGRTLILSTHIDE 1000
Db 1063 SGGNORKLSVAIAVPGGAKVVIDDEPTSGVDPSRRSINDLLKLYRSGRTIIMSTHIDE 1122
Qy 1001 AELIGDRVAVVAGRLCCCGSPFLRRHLGSGYYTLTLVKARLPLTTNEKADTMEGSDVT 1060
Db 1123 ADLLGDIRAIIISQRLYCSGTPLFLKNCFTGTFYTLVR---KMKNIQSQTGCEGCS 1179
Qy 1061 ROEKKNQSGSRVG--TP-----QLLALVOHWVPGARLVEELPHELVLVLPYGAH 1109
Db 1180 ASEGFSTRCPAHIDEITPEQVLDGDNELMDVQHVHPVPEAKLVEICIGOELIFLLPNKFK 1239

Qy	1110	DGSPATLRFRELDTRLAELRLTGYGISTDSLEBEEIPLKVVECAADTDMEDSCGQHCTGI	1167
Db	1240	QRVAJSLFRELEDITLADLGSSFGISDTPLEBEIFLKVT-----EDSDSGFLFAGGT	1290
Qy	1170	AGLDVTLRLK-MPPQETALE-----NG-----EPAGSAPETDQGSGPDVAVGRVO-CWALTRO	1219
Db	1291	OQRENLRHPWLSPREKAROIQPOGSNCCSAEPA-PHEPGOPSSPEARSRLNTGAQLIIQ	1349
Qy	1220	LOQALLIKRFLAARRSRGLFAQIVLPALFVGIALVFSLIVPPPHGVPAIRLSTMYGAQ	1279
Db	1350	HVQALLVKRFHHTIRSHKDFLAQIVLPATFVLALMSIIIVPPEGEYPALILHPMYGOQ	1409
Qy	1280	VSPFSEADPGDPGARLEALLEAOBAG-----LEEPPVOSSHRSFASPEVPAEPAVAK	1329
Db	1410	YTFPSLDQSGEQALAADVLNLKPFGNRCCKEWNLFPEPCGNST-PWKTPSVSPNITH	1468
Qy	1330	VLASGNWTSPSPACOCOSQOGBARLLPDCPAAAGGPPPPQAIVTGSGEVVNLTGRNLSD	1389
Db	1469	LFOKQOWTPEKPSPCRSCTREKLTMLPCEPEGAGGLPPPQRIORSTEILLQDLTNRNISD	1528
Qy	1390	FLVKTYPRLVROGLTKKKWNEVRYGFSIGGRDPGLP-SGOELGRSVSELWALLSPLPG	1448
Db	1529	FLVKTYPALIRSKUSKFPWNEQRYGGISGGKULPIPTGEALVEFLSHLGGIMN-VSG	1587
Qy	1449	GALDR-VLNKLTAWAHSIDAODSLKIWFNNKGHSMTAFVNRRASNATLRAHLPPGPARHA	1507
Db	1588	GPI TREASKEMPAFLKHELETEDNIKWFNNKGHWALVSFLNVAHNAILRTLSHKDKNPKEE	1647
Qy	1508	HSIITLHNPLMLTKEQISEAALMASSVDVLVICSIVPFAMSFVPASPFTVLVIERTVRAKH	1567
Db	1648	YGITVISQPLMLTKEQISEITVLITASDYAVAICVIPAMSFVPASFVLYLIOBRVANKAKH	1707
Qy	1568	LQLMGGLSPTLYMIGNFLMDNCVLYVPACTIWLIFLAFQORAYVAPANLPAKLLLILLYG	1627
Db	1708	LQFVSGVSPITYMLTNFLWDIMNYVASAALVVGIFXGQKATYSPBNLPALTALLMLYG	1767
Qy	1628	WSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLBELFSDQ-KLOBVSRI LKQ	1686
Db	1768	WAVIPMPWPASFLPDVPSTAYVALSCANLFIGINSSTATFTELBENNMVTLFRNAMLRK	1827
Qy	1687	VFLTFPHFCILGRGLIDMVNRQAMADAPERLGDQFQPSPLEWEVVGKNLLAMVTOGPLFULL	1746
Db	1828	LLIIFPHFCILGRGLIDUALSOAVTDVTVARFGBESTNPFQWDLIIGNLVMANAEGVVYLL	1887
Qy	1747	FTLLQHR---SQLLPOPVRVRSPLLGEDEEDVARERERVQOGATQGDVLVLNLTAKVVR	1803
Db	1888	LTLFIQHHFFLTRWVSEPAKE--PII-DEDDDVABERQRIISGNKTDIILNELFKIYS	1944
Qy	1804	GORMPADVRLCLGIPPECCEGLLVNGAGKTKSTFRMWVTGDTLASRGEAVLAGHSVAREPS	1863
Db	1945	GTSSPADVRLCVGRVPGCECFGLLVNGAGKTKTTFKMLTGDVTVTSGDATTAGKSILTNIS	2004
Qy	1864	AHLISMGCYCPQSDAIFELLTGREHLELLARLGRVPEAOVAQTAGSGLARLGLSMYADRPA	1923
Db	2005	DVHQSMGYCQFDVADDLLLTGREHYLYAKLRGVPADEIERVANWSIQSLGLSLYADRLV	2064
Qy	1924	GTYSGGNKRLATALALVGDPAVVFLEPTTGMDDPSARRELWNSLLAVREGSVMLTSH	1983
Db	2065	GTYSGGNKRLSTAIALMGCPPLVLLDEPTTGMDDQARRMLWNTIVSIIREGRAVVLTSH	2124
Qy	1984	SMECECALCSRLATMVNGRFCVLSPOHLKGRFAAGHTLTCLRVPAAIRS-----QPAAAF	2037
Db	2125	SMECECALCRLAIMVNGTFOCGLTIOHLKYKFQDGVIYVIMTKTSKPBEDLLPLDNVPEOF	2184
Qy	2038	VAAEFPGSELREAHGGRRLFPQPPGRCALARYFVELAGHVGAHBGHVEDFSVOTMLEEYF	2097
Db	2185	FQGNFGSVQVRERNYMLQFQV---SSSSSLARIFRLLISHKDSLILLEEYSVTQTTLDQVF	2241
Qy	2098	LYFSKOQGDQED 2109	
Db	2242	VNFAKQQTETHD 2253	

This Page Blank (usp...)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:24:44 ; Search time 27.2021 Seconds
(without alignments)
5119.750 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHHECHFPKPLPSAGT.....QHPKRVQFLDDPSTAETVL 2100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap:*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap:*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap:*
- 5: /cgn2_6/ptodata/1/iaa/ECTUS COMB.pap:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5610.5	51.5	2261	4	US-09-526-193A-1
2	5603.5	51.4	2261	4	US-09-032-438C-118
3	4808	44.1	2273	4	US-09-032-438C-3
4	4730	43.4	2235	3	US-09-032-438C-6
5	3818.5	35.0	1375	3	US-08-665-259-26
6	3818.5	35.0	1375	3	US-08-762-500-26
7	2694.5	24.7	1457	3	US-08-665-259-27
8	2694.5	24.7	1457	3	US-08-762-500-27
9	2693.5	24.7	1472	4	US-09-032-438C-119
10	2507	23.0	1704	4	US-09-032-438C-120
11	2502	23.0	1704	3	US-08-762-500-75
12	2501.5	23.0	1684	3	US-08-665-259-25
13	2501.5	23.0	1684	3	US-08-762-500-25
14	427	3.9	328	4	US-09-724-797-8
15	414	3.8	607	4	US-09-252-991A-18351
16	393.5	3.6	588	4	US-09-489-039A-13579
17	389	3.6	589	4	US-09-328-352-7592
18	375	3.4	594	4	US-09-543-681A-5528
19	374.5	3.4	308	4	US-09-602-787A-348
20	371.5	3.4	335	4	US-09-252-991A-20837
21	363.5	3.3	292	4	US-09-602-787A-352
22	357.5	3.3	271	4	US-09-602-787A-350
23	354.5	3.3	315	4	US-09-328-352-4388
24	350.5	3.2	788	4	US-09-252-991A-28171
25	335	3.1	317	4	US-09-489-039A-10626
26	333.5	3.1	929	4	US-09-252-991A-22946
27	332	3.0	291	4	US-09-107-532A-4205

28	328	3.0	309	4	US-09-252-991A-21204	Sequence 21204, A
29	327	3.0	323	4	US-09-489-039A-12496	Sequence 12496, A
30	319	2.9	922	4	US-09-489-039A-8938	Sequence 8938, Ap
31	316.5	2.9	315	4	US-09-134-000C-6449	Sequence 6449, Ap
32	315.5	2.9	345	4	US-09-252-991A-31957	Sequence 31957, A
33	314.5	2.9	532	4	US-09-543-681A-4646	Sequence 4646, Ap
34	309	2.8	248	4	US-09-710-279-3218	Sequence 3218, Ap
35	306.5	2.8	316	4	US-09-543-681A-6184	Sequence 6184, Ap
36	301	2.8	1280	2	US-08-583-276-19	Sequence 19, Appl
37	300	2.8	1280	4	US-09-767-594-2	Sequence 2, Appli
38	300	2.8	1280	4	US-09-672-810-2	Sequence 2, Appli
39	300	2.8	1280	4	US-09-672-810-5	Sequence 5, Appli
40	300	2.8	1280	4	US-09-672-725C-7	Sequence 7, Appli
41	300	2.8	1280	6	5206352-4	Patent No. 5206352
42	300	2.8	1283	4	US-09-672-810-4	Sequence 4, Appli
43	297	2.7	391	4	US-09-252-991A-20275	Sequence 20275, A
44	297	2.7	1280	2	US-08-752-447-2	Sequence 2, Appli
45	297	2.7	1280	3	US-09-316-167-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-526-193A-1
; Sequence 1, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pinstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526.193A
; CURRENT FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-526-193A-1

Query Match 51.5%; Score 5610.5; DB 4; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;

Qy	1	PPLEHHECHFPKPLPSAGTVPWVQLGICNVNNTCPOLTPGPEPGRUSNFDNLSVRL 60
Db	47	PPYEQHECHFPKAMPAGTLPWVQGIICNANPCFRYPTEGAPGVGVGNFKSIVARLF 106
Qy	61	ADARTVLGGASAHRTLAGLGLKLIATLRAARSTAQ-----DVAEL-----PQ 96
Db	107	SDARLLYSOKDTSMKDMRKVLTLOQIKKSSNKLQDFLVNNETSGFLYHNLSPK 166
Qy	97	PTKOSPLEPPML-----DVAEL----- 113
Db	167	STVDKMLRADVLHKVFLQGVQLHLTSLCNCKSKEEMIQLGDQVSELGCLPREKLAAE 226
Qy	114	-----LTSLRT--ESLGLALGOAQEPHLSLRAEDLAQELIALRLSELVLR---A 159
Db	227	RVLRSNMDILKPIILRTLNSTSPFPFSKELAEATKTLHLSLGTLAQELFSMRSDMRQV 286
Qy	160	LLQPRGTSGLPEL---LSEALCSVRGSSVTGFSLNWYEASDLMLVG---QEPESAL 212

Db 287 PLTNVSSSSTQIYQAVSRIVCHPGGGLKIKSLNWNEDNYKALFGNGTDEEAEATF 346
Qy 213 PDSSLPACSELGALDPSHLRLMLRRLKPLILGKLLFAPDTPPTTRKLAQAQVNRFEEL 272
Db 347 YDNSTTPYCNJDMKNLESSPLSRIIWKALKPLVLGKILYTPDTPATRVQMAEVNKTFOEL 406
Qy 273 TLLADVREVMELGPRIFTFMWDSNSVAMLORLIQMODEGR-RQPRPGGRD----- 322
Db 407 AVFHDLGEMBEELSPKWTTFMENSQEMDLVRMLDSDRNDHFEQQLDGLDWTQAQDIVAF 466
Qy 323 ---HMEALRSPFLDPSGG--YSWODAHADVCHLVGTGRVTECLSLDKLEAAAPSEAAVLS 377
Db 467 LAKHPEDVQS-----SNGSVYTWREAFNEINQAIRTISRFEVCUNLKNLEPIATEWVLIN 521
Qy 378 RALQALLAEHFWAGVFLGPEDDSDPTHEPTDPLGPGHVRIRKIRMDIDVVYTRNKIRDRF 437
Db 522 KSMELLDERKFWAGIVFTGITPGSIELPH-----HVYKIRMDIDNVERTNKIKDGY 573
Qy 438 WDPGPAADPLTDLRYVWGGFVYLODULVERAARVRLSGANPRAGLYLOOMPYPYCVDDVFL 497
Db 574 WDPGRADPEDMYVWGGFAYLODVVEQAIIRVLTGTEKTKGYMQOMPYPYCVDDIFL 633
Qy 498 RVLRSRPLFLTLAWIYSVTLTKVAVREKETRLDTRAMGLSRAVLWLGWFLSCLGPF 557
Db 634 RVMSRSPLEMTLAWIYSVAVIIGIYVEKEARUKETNRIMGLDNLNLSWFSWISSIPL 693
Qy 558 LLSAALVLVLKGLDILPYSHPGVPLFLAAFAVATVYQSFLSAFFSRANLAACGLA 617
Db 694 LVSAGLLVVLKGNLPLPSDPSVVFVFLSVFAVVVTLQCFLLISTLFSRANLAACGGII 753
Qy 618 YFSLYLPLVLCVWRDLRPGGRVAAISLLSPVARGFCESLALLEBOGEGAQHNVTGP 677
Db 754 YFTLYLPLVLCVAMQDYVFTLKFASLLSPVARGFCYFALFEEQIGVQWMDNFESP 813
Qy 678 T-ADVSLAQVSGILLDADALYGLATWYLEAVCPQXGIPPEWNPFRSVCWCPRPKS 736
Db 814 VEEDGFLNTTSVSMULFDTLYGVNTWYIEAVFPQYQIPRPWTFFCTKSYWFGESDEK 873
Qy 737 PAPCFTPLD-PKVLVEEAPGLSPGVSVRSLEKFPSPQPALRGLSLDFYQGHITAFGL 795
Db 874 SHPGSNQKRISIEI CWEEEPHKLGVSIQNLVKYVRDGMKVAVDGLALNFVEGQITSFLG 933
Qy 796 HNGAKTTTILSGLPSPGGSFALICHVDRSSMAAIRHLGVCPQYVNLFDMLTYDEH 855
Db 934 HNGAKTTTWSILTLGTPPTSGTAYILGKOIRSEMSTIRQNLGVCPQHNLFDMLTYVEH 993
Qy 856 WFEYGRKLGLSAAVVPEDRLLODVL-VSKQSVQTRHLSSGMOKLSVAIAFVGGSQV 914
Db 994 IWFYARLKLSEKHVKAEMEQMALDVGLPSSKLKSKTSQISGGMOKLSVALAFVGGSKV 1053
Qy 915 VILDEPTAGVDPASRRGIWELLKYREGRTILILSTHLDLAEALGDRVAVVAGRLCCCG 974
Db 1054 VILDEPTAGVDPYSRGGIWELLKAYRQRTILSTHMDADVLGDRIAIIISHGLCCVG 1113
Qy 975 SPLFLRHLGGYLYTLVKARLPLTTN-----EKADTMGSGVDVDTREKNGS 1022
Db 1114 SSLEFLKNQLGTGYLYTLVKKDVESLSSCRNSSSTSVSLKEDSVSQSSDAGLGSDES 1173
Qy 1023 QGSRVGTPLLALVQHWPVPGARLVEELPHELVLVLPYTGADHGSFATLFRFLDRLAELR 1082
Db 1174 DTLTIDVSAIENLRKGVSEARLVEDIGHETLYVYPYEAKEGAFVELFHEIDRLDSDLG 1233
Qy 1083 LTGYGIDSTSLLEEFLKVEECAADTDMEDGSCGHLCTGIAGLDVTLRLKMPQETALE 1142
Db 1234 ISSYGISSETLLEEFLKVAEESGVDAETSDGTLPARNRRAFG-DKQSLRPFTEDDAAD 1292
Qy 1143 ---NEEPAGSAPETOGSGPDVAG--RVOCWALTQOOLALLKRFILARRSRGLFAQIV 1198
Db 1293 PNDSDIDPESRETDLSSGMDGSGYQVKGWKLTOQQFVALLWKRLLIARRSRKGFPAQIV 1352
Qy 1199 LPALFVGLALVFSLIVPPFGHYPALRISPTWYGAQVSPFSDAPGDPGRALLLEALLQEA 1258

Db 1353 LPAVFCIALVFSLIVPPFGKYPSLELOQPMWYNEQYTFVSDNAPEDGTGLELINALTKDP 1412
Qy 1259 G-----LEBPVQHSSHRFSAPEVPAEVAKVLAGSNWTPESPSPACQCSQFGARR 1308
Db 1413 GFOTRCMEGNIPDPTCQAGEBEWTTAPVQOTIMDLFQNGNWTMNPSPACQCSQKIKK 1472
Qy 1309 LLPDCFAAAGGPPPPQAVTSGSEVQNLGTRNLSDFLVKTPRLVRQGLTKKWNVEVRY 1368
Db 1473 MLPVCPGAGGLPPPPORKQNTADILQDLTGRIISDYLVTYVOI IAKSLKNKTVWNEFVY 1532
Qy 1369 GGFSLG-GRDPGLPSQOELGRSVEELWALLSPUPGGALDRVLKNLTAMAHSLDAQDSLKI 1427
Db 1533 GGFSLGVSNTQALPPSQEVNDALIKQMKHKLAKADSSADRFLNLSLGRFMTGLDTRNNKV 1592
Qy 1428 WFNKQWHSWAPVNRASNAIIRAHLPGPBARHAHSITTLNHPNLNLTQOLSEAAALWASS 1487
Db 1593 WFNKQWHAISPLANINAILRANLQGENPSHYGITAFNHPNLNLTQOLSEVALMTS 1652
Qy 1488 DVVLVSVIVVFAMSFVPASFTLVLIERVTRAKHLQIMGGLSPTLYWLGWFLNLMCNLYV 1547
Db 1653 DVVLVSVIVVFAMSFVPASFTLVLIERVTRAKHLQIMGGLSPTLYWLGWFLNLMCNLYV 1712
Qy 1548 PACIVVLIIFLAFOQRAYVAPANLALLLLLYGWSITPLMYPASFPFSPSTAYVVLTC 1607
Db 1713 PATLVIIIFTCFOQKSVSVSTNLPVALLLLLYGWSITPLMYPASFPFVIPSTAYVVLTS 1772
Qy 1608 INLFIGINGSMATFVLELPSDQKLOEVRILKOVFLIFPHFCLGRGLIDMVRNOAMADAF 1667
Db 1773 VNLFIGINGSVATFVLELFTDNKINNINDILKSFLIFPHFCLGRGLIDMVRNOAMADAL 1832
Qy 1668 ERLGDRQFOSPLRWEVVGVKNLLAMVIOGFLFLFTLLQHRSQLLPQPRVRSPLLGEED 1727
Db 1833 ERGEGNRVSPSLNDVGNLNFAMAVEGVVFFLITVLIOYRFFIRPRPVNAKLSPLENDED 1892
Qy 1728 EDVARERERVQOGATQGVLVLENLTKVYRGQRMPPAVDRCLCLGIPPGCEPGLLGVNGAGK 1787
Db 1893 EDVRRERQRILDGGQNDILEIKELTKIYRRKRPKPAVDRIKVPPECEPGLLGVNGAGK 1952
Qy 1788 TSTFRMVTGDTLASRGEAVLAGHSVAREPSAAHLSGVGCPQSDAIPELLTGRSHLELLAR 1847
Db 1953 SSTFKMLTGDTVTRGDFAFLNKNLSILNTHSVQNNGYCQFQDAITELLTGRSHVEFFAL 2012
Qy 1848 LRGVPEAQVQATAGSLARLGLSWADRPAGTYSGGNKRKLATATALVGDPAVVFDEPT 1907
Db 2013 LRGVPEKEVKCEWAIKGLVKYGEKAGNYSGGNKRKLSTAMALIGGPPVVFDEPT 2072
Qy 1908 TGMDSARREFLWNLAAVVRGSRVMTLSHSMEECEBALCSRLAIMVNGRPRCLGSPHLLK 1967
Db 2073 TGMDSARREFLWNLAAVVRGSRVMTLSHSMEECEBALCSRLAIMVNGRPRCLGSPHLLK 2132
Qy 1968 GRFAAGHTLTLRVPAAARS--QPAAPAAVFAAFPGSELEAHHGRLRQLPGRCALARVF 2025
Db 2133 NREFGDGYTIIVRIAGSNPDLPVQDPFGLAGPGLSVLKEKRNMLQYOL-PSSLSLARIF 2191
Qy 2026 GEELAVHGAHGVDFSVSOTMLEEVFLYFSKDOQKOE---DTEEOKEAGVGVDPAPGLQH 2082
Db 2192 SILSQSKRLHIEDYSVSTQTLQOVFNFAQOQSDDDLKDLSLHKNQTV-VDVAV---- 2246
Qy 2083 PKRVSOFLDDPSTAE 2098
Db 2247 ---LTSFLQDEKVKES 2259

RESULT 2

US-09-032-438C-118
; Sequence 118, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.

APPLICANT: Leppert, Mark
APPLICANT: Dean, Michael
APPLICANT: Singh, Nanda
APPLICANT: Shroyer, No. 6713300h F.
APPLICANT: Smallwood, Philip M.
APPLICANT: Allikmets, Rando
APPLICANT: Lewis, Richard A.
APPLICANT: Li, Yixin
TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
TITLE OF INVENTION: ATP-Binding Cassette Transporter
FILE REFERENCE: BYLR-0065
CURRENT APPLICATION NUMBER: US/09/032,438C
CURRENT FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: US 60/039,388
PRIOR FILING DATE: 1997-02-27
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.2
SEQ ID NO 118
LENGTH: 2261
TYPE: PRT
ORGANISM: Mouse
US-09-032-438C-118

Query Match 51.4%; Score 5603.5; DB 4; Length 2261;

Best Local Similarity 49.6%; Pred. No. 0;

Matches 1118; Conservative 363; Mismatches 577; Indels 195; Gaps 27;

QY 1 PPLEHCHFPNKLPSAGTVPWLOGLICNNVTCPPOLTTGEEPRGLSNFNDLSVRL 60
DB 47 PPYECHCHFPNKAMPSAGTLPWQGIICNANNPCFRYPTCEAPGVGVNFKSVSLF 106
QY 61 ADARTVLGASAHRTLAGLKLATLRAA-----STAQPQ 96
DB 107 SDAQRLLYSORDYSIKDHVKVLMRLROI KHPNSNKLQDFLVNETPSGFLQHNLSLR 166
QY 97 PTKOSPLEPM-----LDVAE 112
DB 167 STVDSLLQKNVLQKVFQYQLHSLASLNGSKLEELIQLGDAEVSALCGLPRKLDAAE 226
QY 113 LL-----TSLLRTESLGLALGOAEPHLSLEAAEDLAELALRLV 155
DB 227 RVLRYNMDILRPVTKLMSSTSLHPTQHLA-----EATVLLDSLGLAQELFSTKWS 279
QY 156 ELR---ALLQPRGTSGLPEL---LSEALCSVRGSSVTGPSLWNYEASDLMELVG--- 205
DB 280 DMQEVMEFLTVNSSSSSTQIYQAVSRIVCGHPEGGGLKI KSLNWIYEDNNYKALFGGNT 339
QY 206 QEPESALPDSLSPACSELIGALDSHPLSLRLRLKPLILGLKLPADPTPTTRKLMQV 265
DB 340 EEDVDFTYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEV 399
QY 266 NRTBEELTLRDREVWEMLPRIPTFNDSNVAMLQRLQMO---DEGRQPRPGGRDH 323
DB 400 NKTQELAVFHDLEGWMELSPOIWTFMENSQEMDLVRLTLLDSRNGDQFWEQKLDGLDWT 459
QY 324 MEALRSFL-----DPGSGYSQMDAHADVGLHVLGTVTECLSLDKLEAAPSEAL 375
DB 460 AQDMAFIAKPNEDVQSNGSVYTWREAFNETNQAIQTSRPMCECNLKLPIPTVRL 519
QY 376 VSRALQLLAEHRFWAGVYFLGPESSDTEPTDGLG---PGHVRKIRKIMDIDVVRINKI 433
DB 520 INKSMELDERKFWAGIVFTG-----ITPDSVELPHVVKYKIRMDIDNVERINKI 569
QY 434 RDRFWDGPAADPLTDLRYVGGFVYLODLVERAAVRVLSGANPAGLYLOMPYPCYVD 493
DB 570 KDGWDPGPRADPFEDMYVYVGGFAYLQDVVVEQAIIRVLTSEKKTGYVYQMPYPCYVD 629
QY 494 DVELRVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRMTWRAGLSRAVLGLWFLSC 553
DB 630 DIFURKRSPLFMTLAWIYSVAVIIKSIYVEKARLUKETWRIMGLDNGILNFSWFVSS 689
QY 554 LGPFLLSAAALLVLKLGDIPLYPHGPVVVFLFLAFAVATVTSQSPLLSAFFSRANLAAAC 613

DB 690 LIPLLVAGLLWILKLGNLPPSDPSVVFVFLSVFAMVTILQCFLISTLFSRANLAAAC 749
QY 614 GGLAYFSLYLYPVLCVARDRLPAGGRVAASLSIPVAFGCESLALLEEGEGAGWNV 673
DB 750 GGIYFTLYLYPVLCVARDYVFSIKIFASLLSPVAFGCGCEYFALFEEQIGVQWDL 809
QY 674 GTRPT-ADVFSLAQVSGILLIDAAALYGLATWYLAECVCGQYGIPEPMFPPRRSVCQPR 732
DB 810 FESPVEEDGNLTAVSNMLFDTFLYGVTWYIEAVFPGQYGIPIPPYFPCCTKSYWFGEE 869
QY 733 PPKSPAPCPTPLD-PKVLVEEAPPGLSPGVSRSLERKFPSPGPOPALRGLSLDFVQHIT 791
DB 870 IDEKSHPGSSQKGVSEICMEBEPTHRLGLVSIQNLVKVYRDGMKVAVDGLALNFVEGQIT 929
QY 792 AFLGHNGAGKTTTLLSGLPPSGGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLT 851
DB 930 SFLGHNGAGKTTTWSIITLGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCPQHNVLDFMLT 989
QY 852 VDEHWFYGRLLGLSAAVVGPEQDRLLQDVGL-VSKQSVQTRHLSGGQMOKLSVAIAFVG 910
DB 990 VEEHWFYARLKLGLSEKHVKAEMEQMALDVGLPPSKLKSQTSLSGGMOKLSVALAFVG 1049
QY 911 GSQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDBAELLGDRVAVVAGRL 970
DB 1050 GSKVILDEPTAGVDPYSRRGIWELLKYRQRTIILSTHMDADILGDIRAIIHSGKL 1109
QY 971 CCCGSPFLRHLGSGYVTLVAKARLPLTTN-----EKADTMESGVDTROEK 1018
DB 1110 CCVGSLLFKNLQGTYYTLVKKDVSSSSSCNNSSTVSLCKKEDSVSSSSDAGLGS 1169
QY 1019 KNGSGSRVGTPOQLLAVHGVPGARLVEELPHELVLVLYTGAHDGFSATLFRLEDTRL 1078
DB 1170 DHESDTLIDVSAISNLIRKHVSEARLVEDIGHETVYVLYPEAAKEGAFVLFHEIDRL 1229
QY 1079 AELRLTGVIGSDTSLEBIFLKVVECAADTMEDG-----SCGQHLCTGI 1123
DB 1230 SDLGSSYVIGSETLEBIFLKVABESGVDAETSDGTLPARNRRAFGDKQC-LHPFTED 1288
QY 1124 AGLDVTLRLKMPPOETALENGEPAGSAPETDOGGSDPAVG--RVQGWALTQQLOALLLK 1181
DB 1289 DAVID-----PNDSID-----PESRETDLLSGMDGKGYQLKGWKLTOQFVALLWK 1335
QY 1182 RFLIARSRRGLFAQIVLPALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSPFSEDA 1241
DB 1336 RLLIARSRRGLFAQIVLPALFVGLALVFSILVPPFGHYPSLELPQWYNEQYTFVSND 1395
QY 1242 PGDPGRARLLLEALQEAQ-----LEPPVQHSRFSAPPEVPAEVAKVASGNWT 1291
DB 1396 PEDMGTOELLNALTQDPGFGTRCMENPIPTDPCLAGBEDWTISFPVQSIVDLFQNGWT 1455
QY 1292 PESPSQACQSQPGARRLLPDCPAAAGPPBPQAVTSGSVEVQNLTCENLSDFLVKTYPR 1351
DB 1456 MKNPSACQSSDKIKKMLPVCPPGAGLPPQKQKTADTLQNLTCENLSDFLVKTYVQ 1515
QY 1352 LVROGLTKKWNVVRVYGGFSLG-GRDPLPSGQELGRSVEELWALLSPLPGCALDRVLK 1410
DB 1516 IIAKSLKNKIWNFRYGGFSLGVNSQALPPSHEVNDIAIKOMKKLLKLTXTDSADRFLS 1575
QY 1411 NLTAWHSLDAQDSLKIWFNNKNGHSMVAFVNRASNAIILRAHLPPGPARHAHSITLNRH 1470
DB 1576 SLGRFMAGLDTKNVWVFNKNGHWAISFLNINNAIILRANLQKGENPSQYGITAFNHP 1635
QY 1471 LNTKEOLSEALMASSVDVLVSI CVVFAMSFVPSFTLVLEBVTYAKHLQMGSLP 1530
DB 1636 LNTKQOLSEALMTTSDVLVSI CVVFAMSFVPSFTLVLEBVTYAKHLQMGSLP 1695
QY 1531 TLYWLGFLDMCNVLPACIVLIFLAFQORA VAVANLPAIILLLILLYGWSITPLMYP 1590
DB 1696 VIYWLNFVWDMCNVVPATLVIIIFGQKQSVSVSTNLVLAALLLLYGWSITPLMYP 1755
QY 1591 ASPFFSVPTAYVYLTCTINLFIGINGSMATFVLEFSDQKLOEVSRLKQVFLIPPHFCL 1650

Db 1756 ASFVKIPSTAYVVLTSVNLFIGINGSVATFVLELFTNNKLNINDILKSVFLIPPHFCL 1815
Qy 1651 GRGLIDWVRNQAMADAFERIGDRQFQSPRLWEVVGKLNLAWVIGQPLFLFTLLQHRSQ 1710
Db 1816 GRGLIDWVRNQAMADALERFGENFVSPFLSNDLVGRNLFAMAVEGVVFFLITVLIOYRFF 1875
Qy 1711 LLPQPRVRSPLLEEDEDVARERERVVQATQGDVLVRLNLTQVYRGQMPADVRLCLG 1770
Db 1876 IRPRPVKAKLPPLNDEDEDVRRQRILDGGQNDILEIKELTKIYRKRKPAVDRI:G 1935
Qy 1771 IPGCEGGLGVNGAGKTSIFRMVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSD 1830
Db 1936 IPGCEGGLGVNGAGKSTTFKMLTGTPTVRGDPAFLNKLSILNIHEVQNMGYCQFD 1995
Qy 1831 AIFELLTGREHELELLARLGVPEAQAQVAGTASGLARLGLSWYADRPAGTYSGGNKRKLIAT 1890
Db 1996 AITELLTGREHVEFPALLRGVPEKEVGFGEWAIRKGLVKGYSKYNSTSGNKRKUST 2055
Qy 1991 ALALVGDPVAVFLDEPTTGMDPSARRFLWNSLLAVVREGSRVMLTSHSMECEALCSRLA 1950
Db 2056 AMALIGPPVAVFLDEPTTGMDPKARRFLWNCALSIVKEGRSVVLTSMSMECEALCTRMA 2115
Qy 1951 IMWNGRFRCLGSPHOLKGRFAAGHTLRLRVPABRS--OPAAAFAAEPFGSELREAHGGR 2008
Db 2116 IMWNGRFRCLGSGVHLKNRFGDGYTVIVRIAGSNPDLPKPVQEPFGLAFPGSVLKEKRN 2175
Qy 2009 LRFOLPGRGRCALARFGEALVHAEGHVEDFSVQTMLEEVFLYFSQDGKDS--DTE 2065
Db 2176 LOYQL-FSSLSLARIIFSILSQSKRLHIEDYSVQTTLDQVFNFAKQSDDDHLKDL 2234
Qy 2066 EQKAGVGVDPAPGLQHPKRVSRQLDDPSTAET 2098
Db 2235 LHKNTQTV-VDVAV-----LTSFLQDEKVKES 2259

RESULT 3

US-09-032-438C-3
; Sequence 3, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 2273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-3

Query Match 44.1%; Score 4808; DB 4; Length 2273;
Best Local Similarity 44.2%; Pred. No. 0;
Matches 1007; Conservative 354; Mismatches 635; Indels 282; Gaps 39;

Qy 1 PPLEHHECHPNKPLPSAGTVPMLOGLICNVNNTCPQLTPGEPGLSNFNSLVSRLL 60
Db 47 PLYSHHECHPNKAMPAGMLPWLQGLFCNVNPPCQSPPTGSPGIVSNNSILARVY 106
Qy 61 ADARTVLGGASAHRTLAG-----LCKLIATLRA----- 88
Db 107 RDPQELLWNPESQHLGRINWTELHILSQFMDTLRTHPERIAGRIRIRIDILKDEETLTLF 166
Qy 89 -----ARSTAQPOFTKQSPLEPPMLDVA----- 111
Db 167 LIKNIGLSDSVYLLINSQVRPEQFAHGVDPDLAKDOIACSEALLERFIIQRRAKTVR 226
Qy 112 -----ELTSLRTESTGLALGQAQEPHLSLEAEED 143
Db 227 YALCSLSQGTLOWIBDTLYANVDFKLFVLPPTLLDSRSQGINL-----RSWGGILSDMSP 282
Qy 144 LAQELLALRSLVEL-----RALLQK--PRGTSGLPELLELSEALCSV-RGPSSTVGPSPNLNWE 196
Db 283 RIQEFTHRPSMQDILLWTRPLMQNGPETFKLMGILSDLLCGYPGEGGSRV-LSFNWYE 341
Qy 197 ASDLMELVGQEPESALP-----DSSLSPACSELICALDPSHLRRLMRLLPLKGLLFA 252
Db 342 DNNYKAPLGDSTRKQDPIYSYDRRTTSCFNALIQSLESNELTKIAPRAAKPMLMGKILYT 401
Qy 253 PDTPTFKLMAQVNRTEELTLRDYREVWEMLGPRIFTFMDNSSNVAMLQRL----- 306
Db 402 PDSFPAARRILKNANSTFEELEHVRKLVKAWEEVGQIWFYFDNSTQNMIRDLTGNPTVK 461
Qy 307 -----QWODREGRQPRGGRDHMEALRSFLDPG-----SGYSMQDAHADVGHVLGT 353
Db 462 DFLNRQLGEGIT-----AEALNLFYKGPRESQADDMANFDWIDFINITDRLRL 512
Qy 354 LGRVTCLSLDKLEAAPSEAAALVSRALQLLAHRFVAGVVLGPESSDTEHPTDLGP 413
Db 513 VNQYLECLVLDKPESYNDETQLTQALSLLEENFWAGVVF-----PDMPWTSSLP 564
Qy 414 GHVRIKIRMDIDVTRTKIRDRFMDGPAADPLTDLRYVYVGGVYLODIVERAAVRLS 473
Db 565 PHVKYKIRMDIDVVEKTKIKDYWSGPRADVEDFRYIWGGFAYLQDMVEQITRSQV 624
Qy 474 GANPRAGLYLQOMPYPYCVDDVFLRSLPLFLTLAWIYSVTLTKAVVRKETRLRD 533
Db 625 QAEAPVGIYLOMPYCFVDDSPMIILNRCFPFMWLAWIYSVMTVKSIVLEKELRKE 684
Qy 534 TMRAMGLSRVWLWGLWFLSCGLPFLSAAALLVLVLKLDILPYSHGQVVFLEAFAVAT 593
Db 685 TLKNQGVSNVAVIWTWFLDSFISMSISFLLTIFIMHGRILHYSDPFILEFLILAFSTAT 744
Qy 594 VTQSFLLSAPFSRANLAAACGLAYSILYPLVLCVAVRDLRDPAGGRVAASLLSPVAPGF 653
Db 745 IMLCFLLSTFPFSKASLAAACSGVIYFTLYLPHILCFPAWQDRMTAEAKKVASLSPVAPGF 804
Qy 654 GCSLALLEBQGGGAQHNWVGTPT-ADVFSLAQVSGLLLLDAAVYGLATWYLEAVCPGQ 712
Db 805 GTYLVRFEEQGLQWSNIGNSFTEGDEFSFLLSQMMLLDAAYVGLLAWYLDQVPGD 864
Qy 713 YGIPEPMNFFRPSYWCGRPPKSPAPCPT-----PL-----DPK-----VLV 750
Db 865 YGTPLPWYFLQESYWLGE-----GCSTREERALEKTEPLETEEDPSHPEGIHDSFF 918
Qy 751 EEAPPGLSQGVSVRSLEKRPFGSPQPALRGLSLDFYQGHITAFIGHNAGKTTTLTSLG 810
Db 919 EREHPGWVPGVCVKNLVKIFEPGCRPAVDRLNITFFYENQITAFIGHNAGKTTTLTSLG 978
Qy 811 LFPSPGGSAPFILGHDRSSMAAIRPHLGVCPQYNVLFMDLTVDEHVMFYGRLLKLSAAV 870
Db 979 LLPPTSGTVLVGGRDIETSLDAVRQSLGCMCPQHNLPHHLTVAEHMLFYAQLKGSQEEA 1038
Qy 871 GPEDRLLDQVGLVSKQSVQTRHLSSGMQRKLSVAIAFVGGSQVILDEPTAGVDPASRR 930
Db 1039 QLEWEAMLEDTGLHHRKNEEAQDLSGMQQRKLSVAIAFVGDAAKVILDEPTSGVDPSRR 1098
Qy 931 GIWELLKLYREGRTLILSTHILDEAELLGDRVAVAGGRILCCCGSPFLFLRHLGSGYLT 990

NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1457 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-665-259-27

Query Match 24.7%; Score 2694.5; DB 3; Length 1457;

Best Local Similarity 40.8%; Pred. No. 4.8e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

QY 750 VEEAPPGLSPGVSRSLEKRFPGSPQALRGSLDFYOGHITAFIIGHNGAGKTTLSILS 809
DB 1 MEEETHPLVVCVDKLVKYNDKLLKLNLYENQVVSFLGHNGAGKTTMSILT 60
QY 810 GLFPSPGSAFLGHDRSSMAIRPHLGVCQYVFLDMLTVDHVMFYGRKLSAAV 869
DB 61 GLFPPTSGSATYIGHDIRTEMDEIRKNLGMCQHNVLFDRLTVEHLLWFYSLKSMQEE 120
QY 870 VGPEODRLLOVGLYSKQSVQTRHLSGCMQKLSVAIAFVGSQVVIIDETAGVDPASR 929
DB 121 IRKETDKMIEDLSNKRHSLVQTLSSGGMKRLSVAIAFVGSRAIILDEPTAGVDPYAR 180
QY 930 RGIWELLKYREGRTLILSTHLEDAELLGDRVAVVAGRLCCGSPFLRRHLGSGYL 989
DB 181 RAIWDLILKYRGRILLSTHMDADLLGDRALIIHSHGLKCCGSPFLKAYDXGRL 240
QY 990 TLVKARLPLTTNEKADTMEGSDVTRQEKNGSQGSRVGTQALLALVQHVWPGARLVEEL 1049
DB 241 TLVKQAPRGTQOEPGLASSPGCPRL-----SSCSEPQVSQFIRKHVASSLLVSDT 292
QY 1050 PHELVLVLPYTGADGSAFATLRELDTRLAELRLTGYSIDTSLEEILKVVEE----- 1103
DB 293 STELSYILPSEAVKKGAFERLFOQLHSLDALHLSGFLMDTTLSEVFLKYSEEDQSLEN 352
QY 1104 CAADT-----DWDGSCCHLCTGTAGLDVTLRLKMPQETALENCEPAGSAPETDOGSG 1158
DB 353 SEADVKEKRDVLPAGEGLTAVGGQAG-NLARCSELAQSASQSASSVGA-RGEETG 410
QY 1159 -----PDVAG-----RVQGWALTQQLOALLKRPL 1184
DB 411 YSDGYDVRPLFDNLQDPDNVSLQSAEAMEALAQVQSGSRKLEGMWLKWRQFHLVKKRFH 470
QY 1185 LARRSRGLFAQIVLPALFVGLALVFLSVLPFPFGHYPALRLSPPTY-----GAQVSFF 1237
DB 471 CARRNSKALCSQILLPAFVVCVAMTVALSVPEIGDLPLPLVLSPSQYHNYTPQPRGNFIYA 530

RESULT 8

US-08-762-500-27

; Sequence 27, Application US/08762500

; Patent No. 6030806

QY 1238 SED-----APGDPGRARILLEALLQAG-----LBEP-----PVQSSH----- 1270
DB 531 NEEROEYRLRLSP-DASPOQLVSTFRLPSGVGATCVLKSPPANGSLGPMNLNLSGSRLLA 589
QY 1271 -RFSAPAEVPAEVAKVLAGSNWTPESPSPA----- 1298
DB 590 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVKXPDEDSLQAWNMSLPPTAGPETWTS 649
QY 1299 -----COCQSGARRLLPDCPAAAGGPPPPQAVTGTSGEVQNLGRNLSDFL 1345
DB 650 APSLRLVHEPVRCTCSAOGTGF-----SCFSSVGGHPPQWRVV-TGDIILTITGHNVSEYL 705
QY 1346 VKTYPRLVRQGLKTKKWNVEVRYGFSGLGRDPGLPS--GOBLGRSVVEELWALLSPLPGG 1403
DB 706 LFTSDRFLH-----RYGATITGNVQKSIPIASFGARVPWPKI----- 744
QY 1404 ALDRVLKNLTAWAHSLSDAQDSLKIWNKGMHSMVAFVNRASNAILRAHLP--GPARH 1460
DB 745 AVRRV-----AQVLYNNKGYHSMPTYLSNNAIILANLPSKSGNPA-- 786
QY 1461 AHSITTLNHLNLTKEQLSEALMASSVDVLVSVICVFAFMSFVPASFTLVLEIERTAK 1520
DB 787 AXITVTNHPMNTKSASLSLYLL-QGTDVVIATPIIIVAMSFVPASFVFLVAEKSTAK 845
QY 1521 HLQLMGGLSPTLYMLGNFMDMNCVLPACIWLIFLAFQORAYVAPANLPALLLLLY 1580
DB 846 HLQFVSGCNVYMLANVMDNLVLPATCCVILFVFDLPAYTSPFNFAVLSFLY 905
QY 1581 GWSITPLMPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELFS-DQKLQEVSRILK 1639
DB 906 GWSITPIMYPASFWEVPSAYVFLVINLFIGITATVATFLLQFEHDKDKLVNYSYLK 965
QY 1640 QVELIFPHFCGLRGILDMVRQAMADAFERLGD-RQFOSPLRWEVVGKLLAMVLOGLPLF 1698
DB 966 SCFLIFPNYNGHGLMAYNEYIYAKIQGDFKMSPFEDWITVTRGLVAMTVGEFVG 1025
QY 1699 LLFTLLQLHRSQLLPQPR---VRSPLLGEDEEDVARERERVVOGATQGDVLVLRNLTKV 1755
DB 1026 FFLITMCQY--NFLRQPOPLPVSTKPV--EDDVVASERQVLRGDANDMDVKIENLTKV 1081
QY 1756 YRGO---RMPAVDRCLGI-PPGSCFGLLVNGAGKTTSTFRMTVGTTLASREAVLAGHS 1811
DB 1082 YKSRKIGRILAVDRCLGVCVPGSCFGLLVNGAGKTTSTFKMLTGDESTTGEAFVNGHS 1141
QY 1812 VARPSAHLGWSVCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLSW 1871
DB 1142 VLKDLQVQOQSLGVCPODPVDELTAHEHLQLYTLRLCIIPWKDEAQVVKWALEKLELTK 1201
QY 1872 YADRPAGTYSNGNKRKLATALLALVGDPAVFLDPTTGMDSARRFLWNSLIIVREGRS 1931
DB 1202 YADRPAGTYSNGNKRKLSTALLGYPAFIFLDPTTGMDSARRFLWNSLIILDKTGRS 1261
QY 1932 VMLTSHSMBECCALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARS-OPAAA 1990
DB 1262 VVLTSMSMBECCALCSRLAIMVNGRFRCLGSIQHLKGRFGDYMITVTKSSQNVKDVVR 1321
QY 1991 FVAEEFPGSELREAHGG-----RLRFOLPPGRCALARVFGELAVHGAHGVDEFSVOT 2045
DB 1322 FPNRNP-----EHAQKTPYKVQYQL-KSEHISLAQVFSKMEQVGVGLGIEDYSVOT 1375
QY 2046 MLEBVLYFSKQKDETEQKEAGVGDVDPAP-----GLOHPKRVSOFL-----DDPSTA 2096
DB 1376 TLDNVFNFAK---KQSDNVQOEAEPSLSPLGLLSLLRPRPAPTELRAVVADEPDL 1432
QY 2097 ET 2098
DB 1433 DT 1434

GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IG5-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1457 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-762-500-27

Query Match 24.7%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 4.8e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 750 VEEAPPGLSPGVRSLEKRPSPQALRGLSLDFYQGHITAFPLHNGAGKTTTLTSL 809
Db 1 MEEETHPLPVVCDVLTQVYKNDKLAANKLSLNLYENQVVSFLHNGAGKTTTMSILT 60

Qy 810 GLFPPSGSAPILGHVRSMAARPHLGVCPQNVFLDMLTVDEHWFVGRGLKLSAAV 869
Db 61 GLFPPSGSATIYGHDIRTEMDEIRKRLGMCQHNVLFDRLTVBEHLWFYSLKMAQEE 120

Qy 870 VGPEQDRLLQDVGLVKSQSVQTRHLSGMQRKLSVAJAFVGSQVWILDEPTAGVDPAS 929
Db 121 IRKETDKMIELELSNKRHSVLQTLSGMKRKLVAJAFVGSRAIILDEPTAGVDPIAR 180

Qy 930 RGIWELLKRYREGRTILSTHLLDEALLGLDRVAVAGRLCCGSPFLFRRHLGSGYL 989
Db 181 RAIWDLILYKPGRTILLSTHMDADLLGDRIALISHGLKCCGSPFLKGYXDGRL 240

Qy 990 TLVKARPLPTNEKADTDMEGSVTRQKNGSGSVGTPPOLLALVQHWVPGARLVEEL 1049
Db 241 TLVKQAPAPGTSQBPGLSSSPGCPRL-----SSCSEFPQVSQFIRKHWASSLLVSDT 292

Qy 1050 PHELVLVLVPTGAGHDSGFATLFRDLTRALBLRTLGTGYSIDTSLDEEIFLKVVEB----- 1103

Db 293 STLSYILPSEAVKKAERLFOOLEHSLDALHSSFLGMDTTLSEVFLKVEDEQSLN 352
Qy 1104 CAADT-----DMEDGSCGHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSG 1158
Db 353 SEADVRESKRDVLPAGELTAVGQAG-NLARCSLAQSAQSASVSSA-RGEBGTG 410
Qy 1159 -----PDVAG-----RVQGWALTRQOLQALLKRFL 1184
Db 411 YSDGYGDRPLFDNLQPDNVSLSQEAEMEAALAVQVQSGSRKLEGWLQMRQPHGLLVKRFH 470
Qy 1185 LABRSRGLFAQIVLPALVGLALVPSLIVPPGHYPALRLSPWY-----GAQVSFF 1237
Db 471 CARNKALCSQILLPAFFVCVAMTVALSVPEIGDPLVLSPSYHNYTPQRGNFPIYA 530
Qy 1238 SED-----APGDPGRARLEALLQEAG-----LEEP-----PVOHSSH----- 1270
Db 531 NEERQEVRLRLSP-DASPPQLVSTFRLPSGVGATCVLKSANGSLGPMNLSSGESRLLA 589
Qy 1271 -RFSAPVPAEVAKVLAGNWTPESSPA----- 1298
Db 590 ARFFDSMCLESFTQGLPLSNFVPPSPAPSDSPVXPDEDSLOANMSLPPTAGPETWTS 649
Qy 1299 -----CQCSQPCGARLLPDCPAAAGPPPPQAVTGSVEVVQNLGTGNLSDFL 1345
Db 650 APSLPLRVLPVHEPVRCTCSAQGTG---SCPSVGVHPQMRVV-TGDLTDTITGHNVSEYL 705
Qy 1346 VKTYPLRVROGLTKKWNVNEVRYGGFSLGGRDPLPS--QGEIERSVEELWALLSPLPGG 1403
Db 706 LFTSDRFLH-----RYGAIITGVNQSIPASFGARVPPMVVKI----- 744
Qy 1404 ALDRVLKULTAWAHSLLDAQSLKIFWNKQHWHSWAFVNRASNAIILRAHLP--GPARGH 1460
Db 745 AVRRV-----AQVLYNNKYHSMPTYLNSLNNAILRANLIPKSKGNPA-- 786
Qy 1461 AHSITTLNHLNLTKEOLSEALMASSVDVLVSI CVVFAFMSFVPASFTLVLIIEBRTAK 1520
Db 787 AYXITVTNHPNKTISASLSLDYLL-QGTDVVIATFIIVAMSFVPASFTLVLEAKSTAK 845
Qy 1521 HLQLMGGLSPTLYWGLNFMCMNYLVPACIVVLIFLAFQORAYVAPANLALLLLLY 1580
Db 846 HLQFVSGCNPVLYWLYNVMNLNLYLPATCCVILFVFDLPAYTSPNTPFAVLSFLLY 905
Qy 1581 GWSITPLMYPASFFSVPTAYVVLTCINLFIGINGSMAFVLELPS-DOKLQEVSRILK 1639
Db 906 GWSITPTMYPASFWFVPSAYVFLIVINLFIGITATVATFLQLFHEHDKDLKVNYSYK 965
Qy 1640 QVFLIPPHFCLGRGLDMVRNOAMADAFERLGD-ROFQSPLRWEVVGKLLAMVIOGPLF 1698
Db 966 SCFLIFPNYNLGHLMEMAYNEIYAKIQGDKMKSPEWDIVTRGLVAMTVEGVG 1025
Qy 1699 LLFTLLQHRSQLLPQR---VRSLLPLGDEEDVARERERVQOGATQGVVLVLRNLTKV 1755
Db 1026 FFLTIMQY--NFLRQORLPVSTKPV--BDDVDVASERQKVLRGDADNDNMVKIENLTKV 1081
Qy 1756 YRGO---RMPAVDRCLGI--PPGECGLGVNGAGKSTFTMTVGTDLASRGEAVLAGHS 1811
Db 1082 YKSRKIGRIILAVDRCLGVCPGECFGLGVNGAGKSTFTKMLTGDESTTGGEAFVNGHS 1141
Qy 1812 VAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSW 1871
Db 1142 VLKDLLOVQOSLGYCQFQFVDPVDELTAHRLQLYTRLCIPKWDQEAQVVKWALEKLELTK 1201
Qy 1872 YADRPAGTYSGNKRKLTATLALVGDPAVVFLDEPTMGDPSARRFLWNSLLAVVRGRS 1931
Db 1202 YADKPAGTYSGNKRKLTATLALVGDPAVVFLDEPTMGDPSARRFLWNSLLAVVRGRS 1261
Qy 1932 VMLTSHMECEALCSFLATMNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARS-OPAAA 1990
Db 1262 VMLTSHMECEALCSFLATMNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARS-OPAAA 1321
Qy 1991 FVAAEFFGSELBRHAGG-----RLRFQLPFGRCALARFVGEALVHGAHGEVDFVSQT 2045

Db 1322 FNRNFP-----EAAHQKTPYKVOYL-KSEHSLAQVFSKMEQVGVGLIEDYSVSQT 1375
Qy 2046 MLEEVFLYFSKQGDDETEOEKAGVGVDPAP-----GLQHPKRVSQFL-----DDPSTA 2096
Db 1376 TLDNVFNFAK---KQSDNVEQOEAPSLSPSLGILLSLLRPPRPAETELRALVADEPEDL 1432
Qy 2097 ET 2098
Db 1433 DT 1434
RESULT 9
US-09-032-438C-119
; Sequence 119, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupeki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 1472
; TYPE: PRT
; ORGANISM: Mouse
; US-09-032-438C-119
Query Match 24.7%; Score 2693.5; DB 4; Length 1472;
Best Local Similarity 40.8%; Pred. No. 6e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;
Qy 750 VEEAPPGLSPGVSVRSLEKRPSPQALRGLSLDFYQGHITAFIGHNGAGKTTTLSTLS 809
Db 16 MEEBPTHLPVVCVDKLTQVYKNDKLAALKSLNLYENVQVSVFLGHNGAGKTTTMSILT 75
Qy 810 GLFPSPGSAFILGHVDRSSMAAIPRHVGCPQNVLPDMLTVDHWFYGRLLKGLSAV 869
Db 76 GLFPPTSGSATYGHDIRTEMDEIRKNLGMCPQHNLPDRLTVEHLWFYSLKSMQEE 135
Qy 870 VGPSQDRLLQDVLGVSKSVQTRHLGGMQRKLSVATAFVGSQVWTLDEPTAGVDPASR 929
Db 136 IRKETDKMIELESLNKRHSVLQTLSGMKRKLVSATAFVGSRAILIDEPAGVDPYAR 195
Qy 930 RGIWELLIKYREGRTLILSTHLLDEAELLDRVAVVAGGRCCGSPFLFRLHGLSGYLL 989
Db 196 RAIWDLILKYKPGRTILLSTHMEADLLGDRIAIISHGKLKCCGSPFLFKGAYKDGRL 255
Qy 990 TLVKARLPLTTNEKADTMEGSDVTRQEKNGSQGSRVGTTPQLIALVQHWVPGARLVEEL 1049
Db 256 TLVKQAPBPGTSQBPGLASSPSGCPRL-----SSCSEFQVSQFIRKXHVASSLLVSdT 307
Qy 1050 PHELVLVLPYTGADHGSFATIFRELDITFLAELRLTGVGIDTSLEEFILKVVEE----- 1103
Db 308 STELSYILPSAVKKGAFERLFQOLESLSLDALHSSGLMDTTLTEEVFLKVSSEEDQSLEN 367

Qy 1104 CAADT-----DMEDGSCQHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSG 1158
Db 368 SEADVKEKRDVLPAGEGLTAVGQAG-NLARCSELAQSQASQASSVGSAR-ROBEGTG 425
Qy 1159 -----PDAVG-----RVOGWALTQOQLALLKRLP 1184
Db 426 YSDGYDYRPLFDNLQDPDVSLOEAEALAQVQGSRLKGLGWLKMRQFHGLLVKRFH 485
Qy 1185 LARRSRGLPAQIVLPALFVGLALVFSLIVPFGHYPALRLSPPTY-----GAQVSFF 1237
Db 486 CARRNSKALCSQIILLPAPFVCVAMTVALSUPEIGDLPFLVLSPSQHYHNYTPRGNFPIYA 545
Qy 1238 SED-----APGDPGRARILLEALQAG-----LEEP-----PVQHSHS----- 1270
Db 546 NEERQEVRLRLSP-DASPOQLVSTFRLPSPGVGATCVLKS PANGSLGPMNLSSGSRLLA 604
Qy 1271 -RPSAPEVPAEVAKVLASGNWTPESPSPA----- 1298
Db 605 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVKPDEDSLOAMNMSLPPTAGPETWTS 664
Qy 1299 -----CQCSPQAGARRLLPDCPAAAGGPPPPQAVTGSQEVVONLTGRNLSDFL 1345
Db 665 APSLPRLVHEPVRTCSAQGTG---SCPSSVGGHPPOKRVV-TGDILTDITGHNVSYL 720
Qy 1346 VKTYPRLVROGLKTKWVNEVRYGFGSLGRDPLGPS--GOELGRSVBEELWALLSPLPGG 1403
Db 721 LFTSDRFLH-----RYGAIITFGNVQKIPASFGARVPMVRKI----- 759
Qy 1404 ALDRVILKNLTAWHSLDAQDSLKITWNNKGWHSWVAFVNRASNAIILRAHLP-----GPARH 1460
Db 760 AVREV-----AQVLYNNKYGHSMPTLYLSNNAIILRAHLRANLPSKGNPAA-- 801
Qy 1461 AHSITTLNHPNLTKEOLESEALMASSVDVLVSCVVPFAMSFVPSFVLVIEBVRTAK 1520
Db 802 AYKITVTNHPNMTKSASLSLDYLL-QGTDVVIAIFIIVAMSFVPSFVFLVAEKSTKAK 860
Qy 1521 HLQLMGGLSPTLYWLGHFLWDMCNLYVPACTIVLIFLAFOQRAYVAPANLPAALLLLLY 1580
Db 861 HLQFVSGCNPIVYLANVWMDMLNLYVPATCCVILFVFDLPAYTSPTNFFAVLSFLLY 920
Qy 1581 GWSITPLMPASFFSVSPSTAYVVLTCINLFIGINGSMAITVLELFS-DQKLQEVSRILK 1639
Db 921 GWSITPIIMPASFWEVFPSSAYVFLIVNLFIGITATVATVLELQLEFHDKOLKVYNSYLK 980
Qy 1640 QVFLIFPHFCGLRGLIDMVRNQAMADAFERLGD-RQFOSPLRWEVVGKLLAMVIOGFLF 1698
Db 981 SCFLIFPNYLNHGLGEMAYNEYNEYAKIQGDKMSPFENDIVTRGLVAMTVGGFVG 1040
Qy 1699 LLFTLLQHRSQLLPQPR---VRSPLLGEDEEDVARERERVQOGATQGDVLVLRNLTKV 1755
Db 1041 FFLTIMQY--NFLRQQRPLPVSTKPV--EDDVVASERQVRVLRGADNDMVKIENLTKV 1096
Qy 1756 YRQO---RMPAVDRCLGI--PPGSCFGLLVNGAGKTTSTFRMTVGTDLTASGEAVLAGHS 1811
Db 1097 YKSRKIGRIILAVDRCLGVCVPGECFGLLVNGAGKTTSTFRMLTGDESTTGEAFVNGHS 1156
Qy 1812 VAREPSAAHSMGVCPOSDAIFELLTGREHLELARLVGPVEAQVQATAGSGLARLGLSW 1871
Db 1157 VLKDLQVQSLGVCPOFDPVDELTAHRLQLYLRICIPWKDEAQVVKWALEKLEUTK 1216
Qy 1872 YADRPACTYSGGNKRLATALALVGDPAVFLDEPTTGMDFPSARRFLWNSLLAVVREGS 1931
Db 1217 YADRPACTYSGGNKRLSTALALIGYPAFLDEPTTGMDFKARRFLWNLILDLIKTGRS 1276
Qy 1932 VMLTSHNECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLTRVPAARS-OPAAA 1990
Db 1277 VMLTSHNECEALCTRLAIWNGRLHCLAGSIQHLKGRFGDGYMITVTKSSQNVKDVVR 1336
Qy 1991 FVAAEFPGSSELRAHGG-----RLRFOLPGRCALARVGEALVHGAHGVDFSVSQT 2045
Db 1337 FFRNRP-----EAAHQKTPYKVOYL-KSEHSLAQVFSKMEQVGVGLIEDYSVSQT 1390
Qy 2046 MLEEVFLYFSKQGDDETEOEKAGVGVDPAP-----GLQHPKRVSQFL-----DDPSTA 2096

```
Db 1391 TLONVFNFAK---KQSDNVEQQAEBSSLPSPGLGSLLSLRPRPAPTELRAALVADEPDL 1447
Qy 2097 ET 2098
Db 1448 DT 1449

RESULT 10
US-09-032-438C-120
; Sequence 120, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h P.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; BEST OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BVLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 120
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-120

Query Match 23.08; Score 2507; DB 4; Length 1704;
Best Local Similarity 33.78; Pred. No. 3.9e-214;
Matches 639; Conservative 314; Mismatches 630; Indels 314; Gaps 47;

Qy 261 LMAQVNRFTBELTLRDVREWMELGPRIFTFMNDSSNAMLQRLQWQDEGRRQPRPG 320
Db 9 LLLKNTVTLQKRVLTVELFL---PLLF-----SGILLWLRKIQSENVNATIYPG- 59
Qy 321 RDHMEALRSFLDPGSGYSWQ----DAHADVGHVLTGLRVTECLSLDKLEAAPSEAAV 376
Db 60 -QSIQELPLFTFPFGDGTWELAVIPSHSDAAKTVTETVRALVINM-RVRGFPSEKDFE 117
Qy 377 SRALQLLAEHFWAGVULFGPDSDDTEHTPTDLGPHGVRIKIRMDIDVTR----- 429
Db 118 DYIRYDNCSSSVLAAVVFEHFNHS---KBPFLPVLAVKYHLRFSTYRRNYMTOTGSFFLK 174
Qy 430 -----TNKIRDRDWDGPAADPLT---DLRYVMGGFVYLQDLVERAAVRVLSGAN--- 476
Db 175 ETEGHTTSLFPLEPNFGP-REPTSPDGGPGYIREGFLAVQHAVDRAIMEYHADAATRO 233
Qy 477 --PRAGLYLOMPYPCVDDVFLRVLSRLPLFTLAWIYSVTLTVKAAVREKETRLRDT 534
Db 234 LFQRLTVTKRFPYPPEFIEDPFLVAIQYQLPLLLLSFTYTALTARIAAVQEKERRLKEY 293
Qy 535 MRANGLRAVLWGLWELSCGLPFLLSAALLVLKLG-----DILPYSHGCVFLFAAF 589
Db 294 MRMMGLSWLHWSAWLFFLLFLIIAASFMTLLFCVKVKKNVAVLSRSDPSLSVLAFLCF 353
Qy 590 AVATVQTSFLISAFPSRANLAAACGLAYFSLYPYVLCVAVWRDLRPAAGRVAASLLSPV 649
Db 354 AISTISFVWSTFFSKANMAAATGGFLYPTIYPIYFFVAPRYNMWMTLSQKLCSCLLSNV 413
```

```
Qy 650 AFGFGCESLALLEEQEGAGQWNVGTRPTA-DVFSLAQVSGILLLLDAALYGLATWYLEAV 708
Db 414 AMANGAQLIGKFEAKGMIQMRDLLSPVNVDDDFCFQGVLMGLLSVLYGLVTWYMEAV 473
Qy 709 CPGQYGIPEPWNFFRRSYWCMGPRPKSPAPCPTPLDPPKLV----EEAPPGLSPGVSV 763
Db 474 FPGQGVQPYFYFIMPSTWCG-KPRAVAGEEDSDPEKALRNEYFEAPEDLVAGIKI 532
Qy 764 RSLEK--RFPFGSPQALRGLSLDFYQGHITAFIGHNGAGKTTTSLSGILSGFPPSGGSAFI 821
Db 533 KHLKSVFRVGNKDRAAVRDLNLNLNLYBGQITVLLGHNGAGKTTTSLMLTGLFPPTSGRAYI 592
Qy 822 LGHDRSSMAAIRPHLGVCPOYVNLFDMLTVDRHWYFGLKGLSAAVGVGPRDRLLODV 881
Db 593 SGYEISQDMVQIRKSLGLCPQHDILFDNLTVAEHLFYAQLKGLSRQKCPCEEVKQMLHII 652
Qy 882 GLVSKOSVOTRHLSSGGMORKLSVAIAFVGSSQVVIDEPTAGVDPAASRRGIWELLKYRE 941
Db 653 GLEDKNRSRFRFLSGGMRKLSIGIALIAGSKVLIIDEPTSGMDAISRRAINWLLQROKS 712
Qy 942 GRTLILSTHLDABELLGDRVAVVAGRCCLCCGSPFLRRHLSGGYYTLTVKARLPITTN 1001
Db 713 DRTIVLTTHPMDEADLLGDRIAIMXGELQCGSSLFKQKYGAGYHMTLVKE----- 765
Qy 1002 EKADTMEGSDTRQEKQSGQSRVGTPOLLALVQHWVPGARLVPELPHLVLVLPYTG 1061
Db 766 -----PHCNPEDISQLVHHHVENATLESSAGAELSFILPRES 802
Qy 1062 AHDSFATLPRELDTRLAELRLGYGIDSITSLSEIEFLKV----- 1100
Db 803 TH--RFEGLFAKLEKKQKELGFIASFGASITTEVEFLRVGKLYDSSMDIOAIQLPALQYQ 860
Qy 1101 VEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGSPD 1160
Db 861 HERRASDWAVDNLCG-----AMDPSDGICALIEBER---T 893
Qy 1161 AVGRVQWALTROOLOALLKRFLLARRSRGIFAQIVLPALFVGLALVSLVPPFGHY 1220
Db 894 AVKLNTGLALHCCQFWAMFLKKAAYSREWKVMAQAQVLPVLTCTVTLALLAINYSSELFDD 953
Qy 1221 PALRLSPTMYGAOVSPFSEDPAGDPGRARLLEALQEALEEPVVOHSSHRFSAPEVPAE 1280
Db 954 PMRLTIGEGRVTVFVSV-----PGTSQLGQQL-----SEH----- 985
Qy 1281 VAKVLASGNWTPSPSPACQCSQPGARRLLPDCPAAAGGPPPPQAVTSGSEVVQNLTRN 1340
Db 986 -----LKDALQAEG--QEPREVILG-----D 1003
Qy 1341 LSDFLVKTYPRVROGLTKKWNVEYRYGSGFSLGGRDPGLPSQOELGRSVEELWALLSPL 1400
Db 1004 LEBFLI-----FRASVEGGGFN----- 1020
Qy 1401 PGCALDRVLKNTAWAHSLLDAQDSLKIWFNNKGWMAFVNRASNAILRAHLPPGPARH 1460
Db 1021 -----RECL--VAASPRDVGERTVNNALFNNQAYHSPATALAVVDNLLFK--LLCGP--H 1069
Qy 1461 AHSITTINHP-----NLNLTKEQLSEALMASSVDVLVSIQVVFAMSVFASFTLVLEER 1515
Db 1070 A-SIVVSNFPQPSALQAQAKDQFNEG---RKGFIDALNL--LFAMAPLASTFSTILAVSER 1123
Qy 1516 VTRAKHLQMLGSLPTLYLWGLNFMCMCNVLPACIVVLIFLAQOQAVVAPANLPALL 1575
Db 1124 AVQAKHVQFSGVHVHSAFWSALLMDLISLIPSLLLLVVFKFADVFARTRDGHMADTLL 1183
Qy 1576 LLLYGWSITPLMPYASFFSFVSTAYVLTVCNLFINGINGSMAFVLEL---FSDOKLQ 1632
Db 1184 LLLYGHWAIIPLMYLNMFFFLGATATRTITIFNLISGI---ATFLMVTIMRIPAVKLE 1239
Qy 1633 EVSRILKQVFLIPHPFCLGRGLDMVRN-----QAMADAFERLGDQFQSP-LRWE 1682
Db 1240 ELSKTLDHVFLVLPNHLGCLGMAVSSFVENYETRYCTSSSEVAAHYCKKNIQYQENFYAWS 1299
```


Db 415 ROLLSPVNVDDFCRQVGLMMLLDVLYGLVWYMEAVFPQGVQVQVFPYFIMPYSYC 474
Qy 730 GPRPKSPAPCTPLDPKVLV-----EEAPPGLSPGVSVRSLEK--RPPGSPQALRGLS 782
Db 475 G-KPRAVAGKEEDSDPEKALRNEVEAFEDLVAGIKIKHLKSVFVGNKDRAAVRDLN 533
Qy 783 LDFYQGHITAFIHNAGAKTTTSLISGLFPPSGGSFAFLGHDRVSSMAAIRPHIGVCPQ 842
Db 534 LNLVEGQITVLIGHNAGAKTTTSLMLTGLFPPTSGRAVYISYEISQDMVQIRKSLGLCPQ 593
Qy 843 YNVLFDMLTVDHVVHVFYGLKGLSAVVGPEODRLLQDVLVSKOSVQTRHLSGGMOKL 902
Db 594 HDILFDNLVFAEHLFYAQLKQKCEPEYQMLHIIGLEDKWNRSRFLSGGKRRKL 653
Qy 903 SVIAFVGGSVQVILDEPTAGVDVPASRRGIWELLKRYREGRTLLSTHLLDEAELLGRV 962
Db 654 SIGIALIAGSKVLLDEPTSGNDATSRRAINDLQKQSDRTIVLTTHFMDAELLGDRI 713
Qy 963 AVVAGRLCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTQBKNGS 1022
Db 714 AIMAKGELQCCGSSFLKQKYGAGYHMTLVKE-----745
Qy 1023 QGSVGTQQLLALVOHVWPGARLVELPHELVLVLPYTCAGDGSFATLPRELDTLAEIR 1082
Db 746 --PHCNPEIDISQLVHHVHPNATLESSAGAEISFILPRESTH--RPEGLFAKLEKKQKELG 801
Qy 1083 LTGYGSDTSLEIFLKV-----VEBCAADTDMEDGSCGOHLCT 1121
Db 802 INSPGASITMEVFLRVGLKVDSSMDTQAIQLPALQYQHERRASDMDVDSNLG-----856
Qy 1122 GIAGLDVTLRLKMPQETALENGEPAQSAPETDQSGDGPVQGVQWALTROQLQALLK 1181
Db 857 -----AMDPSDGIGALIEBER---TAVKLTGLALHCQFQWAMFLK 894
Qy 1182 RELLARRSRGLFAQIVLPALFVGLALVPSLLVPPFGHYPALRLSPTWYGAQVSFSEDA 1241
Db 895 KAAYSWRKWKMAVQAQVPLTCTVLTALLAINYSSELFDPMRLRLTIGEYRTVVPFSV--952
Qy 1242 PGDQGRALLEALQEALEPPVQHSRFSAPVPAEVAKVLASGNWTPSPSPAQC 1301
Db 953 ---PQTSQLGQQL-----SEH-----965
Qy 1302 SQPGARRLLPDCPAAAGPPPPQAVTGVGSEVVQNLTRNLSDPLVKTVPRLVROGLTKK 1361
Db 966 -----LKDALQAEQ--QEPREVLG-----DLEEFLL-----989
Qy 1362 WNEVRYGFSGLGRDPLGQSCQELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSIDA 1421
Db 990 FRASVEGGGFN-----ERCL--VAASFRDVE 1014
Qy 1422 ODSLKIWFNNKGWMAVAVNRASNAILRAHLPGPARGHASTITLHP-----LNLTK 1476
Db 1015 RTVNALFNNQAHSPATALAVDNLKFK--LCCGP--HA-SIVSNFPQPRSAQAOKD 1069
Qy 1477 QLSAALMASSVDLVLSVCVFPAMSPFTVLVLIBERTRAKHQLQMLGSLPTLYWL 1536
Db 1070 QNEG--RKGFDIALNL--LFAFMAFLASTFSLAVSERAVQAKHVQFVGWHVASFWLS 1124
Qy 1537 NFLDMCNVLPACIVLIFLAFQORAVYAPANLPALELLLLLYGWSITPLMYPASFPFS 1596
Db 1125 ALLNDLISFLIPSLLLVVFKAFAFVRAFTRDGHMADTLLLLLYGWAIPMLYMLNPFPL 1184
Qy 1597 VPSTAVVLTCLNLFIGNSGMATVLEL---FSDQKQEVSRILKQVLFIPPHFCICRG 1653
Db 1185 GAATAYTRLTIPNLISGL-----ATFLMTIMRIPAVKUEELSKTLDHVFLVLPNCLGMA 1240
Qy 1654 LIDMVRN-----QAMADAFERLGRQFQSPLE-RWEV--VGKNLLAMWIQG-----P 1696
Db 1241 VSSFVENYETRYCTSSSEVAHCKYNIQYENFVMSAPGVGRFVMSAASGCAYLIL 1300
Qy 1697 LPLFTLLLOHRSQLLPQPRVR-----SLPLGEDEEDVARERVRVQATOG-----D 1745
Db 1301 LFLIETNLLQRLGILCALRRRTLTLYTRMPVL-PEDQDVADERTILAPSPDLLHT 1359

Qy 1746 VLVRLNLTKVYRGQMP--AVDRCLCLGIPPGECFGLLVNGAGKSTSTFRMVTGDTLASRG 1803
Db 1360 PLIIKELSKVYE-QRVPLLAVDRLSLAVKQKGCFCGLGFGNGAGKTTTFFOMLTGBESLTSG 1418
Qy 1804 EAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAOVAQTAGSG 1863
Db 1419 DAFVGGHRISSDVGVRQIRGYCQPDALLDHTGEMLVYARLGRGIPERHIGACVENT 1478
Qy 1864 LARIGLSVADRPACTYSGGNKRKLATALALVGDPAVVPLDEPTTGMOPSAARFLWNSLL 1923
Db 1479 LRGLLEPHANKLVRTYSGGNKRKLSTGLTIGEPAVIFLDEPSTGMDBVARLLMDTVA 1538
Qy 1924 AVVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPA 1983
Db 1539 RARESGKAIITSHSMECEALCTELAIMVQOFKCLGSPQHLKSKFGSGYSLRAKQVSE 1598
Qy 1984 RSQA-----AAFVAABFPFSGSELREAHGRLRQLPQPGRCALARVFGELAVHGAHV 2039
Db 1599 GQOEALBEFKAFVDLTPPGSVLEDBHQGMVHYHL--PGRDLSWAKVFGILEKAKYGVDD 1657
Qy 2040 FVSOTMLEEYFLYFS 2055
Db 1658 YSVSISLEQVFLSPA 1673

RESULT 13

US-08-762-500-25
; Sequence 25, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid

FILE REFERENCE: 2653-40
CURRENT APPLICATION NUMBER: US/09/724,797
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 60/111,325
NUMBER OF SEQ ID NOS: 95
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 328
TYPE: PRT
ORGANISM: Bacteria
US-09-724-797-8

Query Match 3.9%; Score 427; DB 4; Length 328;
Best Local Similarity 34.9%; Pred. No. 3.2e-29;
Matches 112; Conservative 59; Mismatches 142; Indels 8; Gaps 5;

QY 1750 RNLTKYRGQRPDAVRLCLGIPPECFCGLLVGNAGKSTFRMTVGTDLASRGAIVLAG 1809
DB 12 RDLKVFVGQTR--AVDGLDLVVRAGTIHGVLPNGAGKTTAIKMLATLMRPTSGTASVLG 69

QY 1810 HSAVREPSAAHLSMVCPOSDAIFELLTCREHLELLARLVPEAQAQTAGSGLARLGL 1869
DB 70 HDVREAAEVRRIIGLTGTMVSDEMTGVQNLILAGRLQGLRHASAAAEQLMEAFDL 129

QY 1870 SWYADRPAGTYSGGNKRKLATLALVGPVAVVFLDEPTTGMDSARRFLWNSLLAVVREG 1929
DB 130 TEVGRVLKVTSGGQRRIDVAASVVTPELLFLDEPTTGLDPRSEWEMIRALVRDG 189

QY 1930 RSVMLTSHMECEALCSRLAIMVNGFRCLGSPQHLKGRFAAG--HTLTLRVPAARSQPA 1988
DB 190 GTVLLTTQVLDADHLADELTLDHGRIVAQGTTPPELKASRAAGVLDVRLRDPERRADAG 249

QY 1989 AAFVAEPFSGELREAHGRLRFQLPFGRCALARVFGELAVHGAHGVDEFSVOTMLE 2048
DB 250 ALLAKAVGAADL--DSDPARLSVRVTDPDRAALA--LGELARAGIH--VDDFTLQGPSLD 304

QY 2049 EVFLYFSKDGKDETEEQE 2069
DB 305 TVFLALTGHSTVDASEEEAE 325

RESULT 15
US-09-252-991A-18351
Sequence 18351, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18351
LENGTH: 607
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18351

Query Match 3.8%; Score 414; DB 4; Length 607;
Best Local Similarity 17.7%; Pred. No. 1.4e-27;
Matches 223; Conservative 100; Mismatches 235; Indels 702; Gaps 26;

QY 732 RPPKSPAP-----CPTPLDPKVLVEEAPPLSPGVSVRSLEKRPFGSPQPALRGSLDFY 786
DB 3 RSMRPAPVSGSVEMQVSPAIAAVPAGADEAAVIEDVKHF--GDVKALRGLSARIH 60

QY 787 QGHITAFILGHNGAGKTTTTLISLGLFPPSPGSAFLIGHDVRSSMAAIRPHLGVCQYNVL 846

DB 61 YGRLTGLVGPDPGAGKTTLMRILTLGLVNPAGRVTLTLAGYDVVKNDAIHVASGMPQRFGL 120
QY 847 FDLMTVDDEHVMFYGRKLGLSA-----AVGPEQDRLLQDVLGVLKQSVQTRH 893
DB 121 YEDLSVMENMRLYAQLRGMDADRNAELFAELLDFTRLGPFPTKRL-----AGK 167

QY 894 LSGGMORKLSVAIAFVGGSQVVILDEPTAGVDVPASRRGIWELLKYR--EGRTLILSTHLL 952
DB 168 LSGGMKQKLGACALMARPKVLLLDPEGVGVDVPSRQDLWRMVQALTDGEMAVVWSTAYL 227

QY 953 DEAEELLGDRVAVVAGGRLCCCGSPFLRRRLHSGSYLLTLVKARLPLTTNEKADTMEGSV 1012
DB 228 DEAEER-----CESVLLNQ----- 241

QY 1013 DTRQEKNGSQSRVGTTPQLLALVQHVWPGARLVPELPHLVLPYTCAGHDGSPATLPR 1072

DB 242 ----- 241

QY 1073 ELDTRLAELRLTYGISTLEEFKLVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRL 1132

DB 242 -----GQLLFDG----- 248

QY 1133 KMPQETALENGEPAGSAPETDQSGPDVAGRVQGMALTROQLQALLKRFLLARRSRG 1192
DB 249 --PQGE-----LTAQ-----LEGRS--- 261

QY 1193 LPAQIVLPALFVGLALVFSLIVPPPGHYPALRLSPTMYGAQVSFFSEDAPEGDPGARLLE 1252

DB 262 -----FRLENVGAERRAVLTE 277

QY 1253 ALLQEBAGLEPPVQHSRFSAPVPAEVAKVSLASGNWTPSPSPACQSQPGARRLLPD 1312

DB 278 AL-----D 280

QY 1313 CPAAGGPPPPQAVTSGEVVQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFS 1372
DB 281 LPSVSDG-----VIQAGVRV-----VLREGAPTEQ----- 306

QY 1373 LGGRDPGLPGQELGRSVEELWALLSLPLPGNALDRVLKNLTAWAHSLEDAQDSLKTFWNK 1432

DB 307 ----- 306

QY 1433 GWHSMVAFVNRASNAIHLAHLPPGPAPARHAHSITTLNHLNLTKEQLSEAAALMASSVDVLV 1492

DB 307 -----IQALADRA-----QVQLAPVPAR----- 324

QY 1493 SICVVFAMSFVPASFTLVLIEERVTRAKHLQMLGGLSPTLYWLNFLWDMCMNYLPACIV 1552

DB 325 -----FEDAFI-----DLGG----- 335

QY 1553 VLIPLAFOORAVVAPANLPAALLLLLLLYGWSITPLMYSPASFFSVPTAYVVLTCLNLF 1612

DB 336 ----- 335

QY 1613 GINGSMATFVLELFSQKLQBSRILKQVFLIPHFCLGRGLIDMVRNQAMADAPERLGD 1672

DB 336 ---GPGGTSLA-----ERL--- 347

QY 1673 RQFOSPLRWEVVGKLLAMVIOGFLPLFTLLQHRSQLPQPRVRSPLLGEEDEDVAR 1732
DB 348 ---SPV-----ELGSDVA- 357

QY 1733 ERERVQCATQGDVLVLRNLTKVYRGQRPDAVRLCLGIPPECFCGLLVGNAGKSTFR 1792

DB 358 -----VSCNRLTK--RFGEFTATDQVSFEVQKEIFGLLGNAGKSTTFK 401

QY 1793 MVTGDTLASGEAVLAGHSVAREPSAAHLSMVCPOSDAIFELLTCREHLELLARLVGP 1852

DB 402 MLCGLLKPTAGEAHVGHDLRHATGAQSOLGYMAQKFSLYGLLSVRQNLRFESAGVYGLE 461

QY 1853 EAQVATAGSGLARLGLSWYADRPAGTYSGGNKRKLATLALVGPVAVVFLDEPTTGM 1912

Db 462 GNVRRERIEEMIAFDLGDWLSATPDSLPGLGHKQRLALACSLMHRPPVLFDEPTSGVDP 521
Qy 1913 SARFRLNLSLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAA 1972
Db 522 ITRREFWTHINGLARKGVTIMVTHFMDEAE-YCDRVAMLSRARLIALDTPDALK-RIAA 579

Search completed: December 29, 2004, 22:54:05
Job time : 43.2021 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 47.9746 Seconds
(without alignments)
4211.715 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLFHECHFPKPLPSAGT.....QHPKRVQFLDDPSTAEATVL 2100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5531.5	50.8	2201	2 A54774	ATP binding casset
2	2863.5	26.3	1529	2 A59189	ATP-binding casset
3	2694.5	24.7	1472	2 B54774	probable ATP-bindi
4	2503	23.0	1704	2 S71363	ATP-binding casset
5	2502	23.0	1704	2 A59188	probable ABC-type
6	1829	16.8	1802	2 T33783	hypothetical prote
7	1813	16.6	1816	2 A84845	probable ABC trans
8	1603	14.7	1317	2 C88925	protein F33E11.4 [
9	1600.5	14.7	1447	2 T15200	hypothetical prote
10	1408.5	12.9	1758	2 F88559	protein c48B4.4b [
11	1407	12.9	1767	2 S60124	transport protein
12	1405	12.9	1704	2 T42749	ATP-binding casset
13	1376	12.6	1246	2 T00826	hypothetical prote
14	1345	12.3	269	2 T46467	hypothetical prote
15	1038.5	9.5	1564	2 T27121	hypothetical prote
16	937.5	8.6	1431	2 T22748	hypothetical prote
17	901.5	8.3	373	2 T47150	hypothetical prote
18	741.5	6.8	1011	2 T07712	probable ABC-type
19	665	6.1	895	2 T07714	probable ABC-type
20	650	6.0	900	2 T07717	probable ABC-type
21	625.5	5.7	722	2 T07716	probable ABC-type
22	618.5	5.7	925	2 T07713	probable ABC-type
23	539	4.9	1336	2 T18288	ABC transport prot
24	435	4.0	582	2 H95950	probable ABC trans
25	429	3.9	330	2 S27707	daunorubicin resis
26	423.5	3.9	664	2 T07715	probable ABC-type
27	419.5	3.9	342	2 T36505	probable ABC-type
28	410	3.8	339	2 A52102	ABC transporter AT
29	408.5	3.7	333	2 D72492	probable ABC trans

30	408.5	3.7	339	2 S74048	probable daunorubi
31	403.5	3.7	326	2 JN0849	hypothetical 35K p
32	403	3.7	332	2 T36741	probable ABC-type
33	402	3.7	347	2 S76278	ABC-type transport
34	401.5	3.7	583	1 B64816	ABC-type transport
35	398.5	3.7	583	2 H90737	hypothetical prote
36	395.5	3.6	583	2 A85588	hypothetical prote
37	393.5	3.6	308	2 D97589	probable ATP-bindi
38	393.5	3.6	308	2 A12810	hypothetical prote
39	393	3.6	254	2 A11554	daunorubicin resis
40	393	3.6	325	2 S32908	hypothetical prote
41	392	3.6	196	2 T12512	hypothetical prote
42	390	3.6	314	2 D97318	ABC-type MDR trans
43	387.5	3.6	246	2 S75436	hypothetical prote
44	387.5	3.6	331	2 A52008	ATP-binding protei
45	386.5	3.5	257	2 C84147	ABC transporter (A

ALIGNMENTS

RESULT 1

AS4774
ATP binding cassette transporter ABC1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 02-Feb-2001
C:Accession: A54774
R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A:Reference number: A54774; MUID:94375008; PMID:8088782
A:Accession: A54774
A:Molecule type: mRNA
A:Residues: 1-2201 <LUC>
A:Cross-references: GB:X75926; NID:9495256; PIDN:CAA5330.1; PID:9495257
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop
P:856-1047/Domain: ATP-binding cassette homology <ABC1>
F:873-880/Region: nucleotide-binding motif A (P-loop)
F:1869-2060/Domain: ATP-binding cassette homology <ABC2>
F:1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match	50.8%	Score	5531.5	DB 2	Length	2201			
Best Local Similarity	49.4%	Pred.	No. 0						
Matches	1107	Conservative	363	Mismatches	574	Indels	195	Gaps	27
QY	15	LPSAGTVFWLQGLICNVNNTCFPOLTPGEEPRGLSNDSLVSRLLADARTVLGGASAH	74						
Db	1	MPAGTLPWVQGIICNANNPCFRYPTEGEAPGVGNFNKSIVSRFLSDAQRLLLYSQDRT	60						
QY	75	TLAGLGKLIATRAAR	107						
Db	61	SIKDMHKVRLRLQIKHPNSNLKLDPLVDNETFSGFLQHNLSLPRSTVDSLLQXNVGLQ	120						
QY	108		114						
Db	121	KVFLQGYQLHLASLNGSKLEETIQLGDAEVSALCGLPRKKLDAERVLRYNMDILKPVV	180						
QY	115	-----TSLRSTESLGLALGOAQEPHLSLLEAAEDLAQELLALRSVLVEL	166						
Db	181	TKLNSHSLPTQHLA-----EATTVLLDSLGLLAQELFSTKSWSDMRQEVMLFTVNS	233						
QY	167	TSGPLEL-----LSEALCSVRGSSSTVGSPLNWYEASDLMLVG	219						
Db	234	SSSSTQIYQAVSRIVCGHPGGGLKIKSLNWYEDNNYKALFGGNNTEEDVDFTYDNSTTP	293						
QY	220	ACSELIGALDSHPLSRLLRRLKLLIAGKLLFADPTPTFRKLMAQVNRTEELTLRDVR	279						
Db	294	YCNDLMKNLESSPLSRILIKALKPLLVGKILYTPDTPATRQVMAEVNKTQFELAVFHDLE	353						
QY	280	EVWEMLGPRIFTFMDSSNVAMLRLLQMQ--DEGRRQPRPGGRDHMBALRSFL	331						
Db	354	GMWEELSPLQITWFMEANSQEMDLVRLTLDSDRGNDQFWEQKLDGLDWTQADIMAFKLNKPED	413						

```
QY 332 --DPGSGGYQWQADAHVDVGLVGTGRVTECLSLDKLEAAPSEAAVLSRALQLLAEHRFW 389
Db 414 VQSPNGSVYTWREAFNETNQAIQITISFMCEVNLKLEPIPTVRLINKSMELLDERKFW 473
QY 390 AGVVFLGPEDSSDTEPTHTDGLG--PGHVRIRKIRMDIDVYTRTNKIRDRFDWPGPAADPL 447
Db 474 AGIVFTG-----ITPDSVELPHVVKYKIRMDIDNVERINKIKDKGYWDGPRADPF 523
QY 448 TDURYVWGGFYLODLVERAAVRVLSGANPRAGLYLOOMPYPYPCVVDVDFVLVLSRSLPLF 507
Db 524 EDMRYVWGGFYAYLDVVEQAIIRVLTSEKKTGYVVOQMPYPCVVDVDFILRVMSRSMPLF 583
QY 508 LTLAWIYSVTLTKVAVREKETRLDTRWAMGLSRAVLWGLWFLSCGLPGLLSAALLVLY 567
Db 584 MTLAWIYSAVITKISIVYEKEARUKETMRIMGLDNGILWFSWFSSLIPLILVSAGLLVVI 643
QY 568 LKLGDIIPYHPGPFVFLFAAFVATVTSQFLLSAFPSRANLAAACGLAYFSLPYLYVL 627
Db 644 LKLGNLPLPYSDPSVVFVLSVFAMVTILQCFLISTLFSRANLAAACGIIYFTLYLYVL 703
QY 628 CVAWRDLRPGGRVAASLLSPVAFPGFCESLALLEBQEGAGAHNVGTRPT-ADVFLSLAQ 686
Db 704 CVAWQDVYVFSIKIFASLLSPVAFPGFCYFALFEQIGVQMDNLPESPVEEDGFNLTT 763
QY 687 VSGLLLDAAVGLATWVLEAVCPGOYGIPEPNWPFERRSVWCGPRPKSPAPCPTPLD- 745
Db 764 AVSMMLTDTLYGVMTWYIEAVFPQYGIIPRPTFPCTKYWFGEEDIEKSHPGSSQKGV 823
QY 746 PKVLVEAPPGLSPGVSRSLERKFPGPSOPALRGLSLDFYQGHITAFILGHNGAGKTTTL 805
Db 824 SEICMEEPHLRLGVSIQNLVKYRDGMKVAVDGLALNFYEQGITSFLGHNGAGKTTM 883
QY 806 SILSGLPPPGGSAFILGHVDRSMAAIRPHLGVCPQYNVLFDMLTVDENHWFYGRUKGL 865
Db 884 SILTGLPPPTSGTAYILGKDIRSEMSSIRQNLVCPQHNVLFDMLTVEEHIWIFYARUKGL 943
QY 866 SAAVGPEDRLLODVLG-VSKQSVOTHLISGGNORKLSVAIAFVGGSOVVLDEPTAGV 924
Db 944 SEKHVAEMEQMALDVLPLSPKLSKTSQSLGGNORKLSVALAFVGGSKVVLDEPTAGV 1003
QY 925 DPASRRIGIWELLKYRREGRTLIITSHLDLAEALIGDRVAVVAGGRCLCCGSPFLRLRHLG 984
Db 1004 DPFYRRIGIWELLKYRGRITLITSHLDEADILGDRALIIHSGKLCVGSLSFLKNQGL 1063
QY 985 SGYYTLVKARLPJTTN-----EKADTDMEGSVDTTQEKKNKGSGRVRGTPQL 1032
Db 1064 TGYTLVLVKDVESLSRNSSTVSCLKKEDSVSQSSDAGLSDHESDITLIDVSAI 1123
QY 1033 LALVOHWVPGARLYVEELPHELVLVLPYTGADHGSFATLRELDTRLAELRLTGIGISDTS 1092
Db 1124 SNLRKHVSARLVEDIGHLETVLPLVEAKGAFVLFHEIDRLSDGLGSISSYISETT 1183
QY 1093 LEEIFLKVBECAADTDWEDG-----SCQHLCTGIAGLDVTLRLKMPQ 1137
Db 1184 LEEIFLKVABESGVDAETSDGTLPARNRARFGDKQSC-LHPFTEDDAVD-----PN 1234
QY 1138 ETALENGEPAGSAPETDGGSPDAVG--RVQGWALTQQIQOALLKRLFLIARRSRRLGFA 1195
Db 1235 DSDID-----PESRETDLGMDGKGYLKGWMLTQQQFVALLWKRLLIARRSRKGFFA 1289
QY 1196 QIVLPALFVGLAVFSLIVPFGHYPALRLSPMYGAOVFFSDAPGDCPRALLLEALL 1255
Db 1290 QIVLPFAVFCIALVFSILVFPFGYPSLEIQFWMYNEQYTFVSNADPEDMGTOELLNALT 1349
QY 1256 QEAG-----LEPPPVQHSRHSFAPEVAEVAKVLASGNWTPESPSPACQSQPG 1305
Db 1350 XDPFGFTRCMEGNIPDTPCLAGEDWTISFPQSIIVDLFQNGWNTWKNFSPACQSSDK 1409
QY 1306 ARLLLPDCAAGAPPPQAVTGSGEVQNLTGRLNSDFLVKTYVRLVROGLKTKYKWNE 1365
Db 1410 IKKMLPVCPPGAGGLPPQPKQTADILQNLTGRLNSDFLVKTYVQIIIAKSLKNKIWNE 1469
```

```
QY 1366 VRYGGFSLG-GRDPLGSLGQELGRSEVEELWALLSPLPGGALDRVLKMLTAWAHSLDAQDS 1424
Db 1470 FRYGGSFSLGVSNSQALPPPSHEVNDALIKOMKLLKLTKDTSADRFSLSGRFMAGLDTKNN 1529
QY 1425 LKTIWFNNKGWHSVAVFNRAASNAILRAHLPPGPARHAHSITTLNHPNLNLTKQOLSEALM 1484
Db 1530 VKVWFNNKGWHAISFLNINNAILRANLQGENPQYGITAFNHPNLNLTKQOLSEVALM 1589
QY 1485 ASSVDVLVSICTVVPAMSFVPSFTLVLIERVTRAKHLQLMGGLSPTLYWGLHFLWDMCN 1544
Db 1590 TTSVDVLVSICTVVPAMSFVPSFTLVLIERVTRAKHLQFISGVKVPVYWLSPVWDMCN 1649
QY 1545 YLVPACTVLVILFLAFOORAVVAPANLPALILLLLGLWSTPLMYPASFFSPSTAYVV 1604
Db 1650 YVVPATLVIIIFTCFQKQSVSVSTNLPVLAALLLLGLWSTPLMYPASFFVKIPSTAYVV 1709
QY 1605 LTCINLFIGINGSMATFVLELFSQDKLQEVSRILKQVLFIPPHFCLGRGILDMVRNQAMA 1664
Db 1710 LTSVNLFIGINGSVATFVLELFTNNKLNINDILKSVFLIFPHFCLGRGILDMVRNQAMA 1769
QY 1665 DAFERLGDROFQSPLEWVGVKNLLAMVIOGPLFLLFTLLQHRSQLLPQPRVRSPLLG 1724
Db 1770 DALERFGENFVSPSLWDLVGRNLFAMAVEGVVFFLTIVLIQYRFFIRPRPVKAKLPLN 1829
QY 1725 EDEDVARERVRVQATQGDVLVLRNLTKVYRGQRMPAVDRLCLGIPPGCECFLGVLNG 1784
Db 1830 DEDEDVRRERQRIIDGGQNDILEIKELTKIYRKRKPAVDRLCIGIPPGCECFLGVLNG 1889
QY 1785 AGKTSFRMVVTGDTLASRBAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLEL 1844
Db 1890 AGKSTTFKMLTGTPTVRGDAFLNKNLSILNIEHVHQNMGYCPQSDAITELLTGREHVEF 1949
QY 1845 LARLGVPEAQVTAQTAGSLARLGLSWYADRPACTYSGGNKRKLATALALVGDPAVVFLD 1904
Db 1950 FALLRGVPEKEVGKFGWEAIRKLGLVKYGEKYSNYSGGNKRKLSTAMALIGGPPVVFLD 2009
QY 1905 EPTTGMDPSARRFLMNSLLAVVREGSVMLTSHSMEECEALCSRALMVGNRFCRLGSPQ 1964
Db 2010 EPTTGMDPKARRFLMNCALSIVKEGSRVLTSHSMEECEALCTFMALMVGNRFCRLGSPQ 2069
QY 1965 HLKGRFAAGHTLRLVPAARS--OPAAAFVAAFPFGSELREAHGGRRLRFOLPPGGRCALA 2022
Db 2070 HLKNRFGDGYTVIVRIAGSNPDLKPVQEFFGLAPFGSVLKEKERNMLQYL-PSSLSLA 2128
QY 2023 RVFGLAVHGAEGHVEDFSVOTMLEVFLYFSKQDGKDE--DTEQKEAGVGVDPAPG 2079
Db 2129 RIFSILSQSKXRLHIEDYSVSTLTDQVFNFAKQSDDDHLKDLSLHKQTV-VDVAV- 2186
QY 2080 LQHPKRVQSFOLDPSTAET 2098
Db 2187 -----LTSFLQDEKVKES 2199
RESULT 2
A59189
ATP-binding cassette transporter - human (fragment)
N/Alternate names: KIAA1062 protein
C/Species: Homo sapiens (man)
C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C/Accession: A59189
R/Kikuno, R.; Nagase, T.; Ishikawa, K.; Hiroseawa, M.; Miyajima, N.; Tanaka, A.; Kotani,
DNA Res. 6: 197-205, 1999
A/Title: Prediction of the coding sequences of unidentified human genes. XIV. The comple
A/Reference number: Z22961; MUID:99397452; PMID:10470851
A/Accession: A59189
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-1529 <KIK>
A/Cross-references: UNIPROT:Q9HC28; GB:AB028985; NID:95689460; PIDN:BAAB3014.1; PID:d104
A/Experimental source: chromosome 9; clone nj03579; clone lib pluescriptII SK plus; tis
C/Genetics:
A/Map position: 9
A/Note: KIAA1062
```

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

```
Query Match      26.3%; Score 2863.5; DB 2; Length 1529;
Best Local Similarity 41.1%; Pred. No. 3.4e-171;
Matches 644; Conservative 246; Mismatches 438; Indels 239; Gaps 36;

Qy 699 GLATWLEAVCGQVGIPEPNPFRSVMCGP-RPPKSPAPCPPLDPKVLV-----750
Db 1 GILTWIEAVHVGMYGLPRPMPYPLQKSYWLSGRTEAWEMSWPWPARTPLRSMBEEDQAC 60

Qy 751 -----EEAPECLSPGVSVRSLEKPEPSPQALRGLSLDFYQGHITAFIAGHN 797
Db 61 AMESRRFBETRGMEETHPLPLVVCVDKLTQVYKDDKLLANKLSLNLYENQVVSFLGHN 120

Qy 798 GAGKTTTILSLGPPPPSGGSFALIGHDVRSSMAAIRPHLGVCPQVNVLFDMLTVDHVVW 857
Db 121 GAGKTTTMSILTGLPPTSGSATYGHDIRTEWDEIRKNLGNCPQHNVLFDRLTVEHLW 180

Qy 858 FYGRUKGLSAAVGPEDQRLLQDGLVSKQSVOTRHLSGGMQRKLSVAIAFVGGQVIL 917
Db 181 FYSRLKSWAQEBIRREMDKMIEDLELSNKRHSLVQTLGGMKRKLSVAIAFVGGSRAIL 240

Qy 918 DEPTAGVDPASRGITWELLKYREGRTLLSTHHLDEALLGDRVANVAGGRLCCGGSL 977
Db 241 DEPTAGVDPYARRAIWDILKYKQRTILLSTHMDADLLGDRITAIISHGKLKCCGSL 300

Qy 978 FLRRHLGSGYILTVKARLPLTNEKADTDMGSDVTRQEKNGSQGSRVGPQLLALVQ 1037
Db 301 FLKTYGQGYRLTVKR--PAEPGQBPGLASSPPGRAPLSSCE-----LQVSQIR 352

Qy 1038 HWVPGARLVEELPHLVLVLPVTYGAHDSFATLPREDLTRLABRLTYGIGSDTSLERIF 1097
Db 353 KHVASCLLVSDTSTSLVILPEAAKKAFAERLFQHLERSLDALHLSFGLMDTLLEVF 412

Qy 1098 LKVVEE-----CAADT-----DMEDGCGQLCTGAGLDTVLRLKMPQETALENGEP 1146
Db 413 LKVSEEDQSLNSEADVKESKOVLPFGAEPASGEGHAG-NLARCSLTSQASLQSASS 471

Qy 1147 AGSAPETDQGG-----PDAVG--RYQGWALTR-----1172
Db 472 VGSA-RGDEGAGTYDVGYRPLFDNPQDPDNVSLQVEAEALSRVQGSRLDGGWLKV 530

Qy 1173 QOLQALLKRFLLARRSRGLPAQIVLPALFVGLALVSLIVPPGHPYALRLSPMY--1230
Db 531 RQFHGLLVKRFHRCARRSKALFQILLPAFFVCVAMTVALSVEIGDPLPLVLSPSQVHN 590

Qy 1231 -----GAQVSFSED-----APGDPGRALLLEALIQEAG-----LEEP-----PV 1265
Db 591 YTPRGNFPIPYANERREYRLRLSP-DASPOOLVSTFRLPSGVGATCVLKSPANGSLGPT 649

Qy 1266 QHSSH-----RPSAEVPAEVAKVASGNWTPESPSPA-----1298
Db 650 LNLSGESRLLAARFDSKCLSFQGLPLSNFVPPPPSPAPSDSPASPDDEDLQWNVSL 709

Qy 1299 -----CQCSQOPGARLLPDCPAAAGPPPPPPQAVTSGSEYVQ 1334
Db 710 PPTAGPEMTSAPSLPRLVREPVRTCSEAQGTG---SCPSSVGGHPQPMRVV-TGDIILT 765

Qy 1335 NLTGENLSDFLVKTYPRLVROGLTKKWNNEVRYGFSGLGRDPLGSGELGRSVEELW 1394
Db 766 DITGHNVSEYLLFTSDRFLH-----RYGAITFG-----794

Qy 1395 ALLSPLPGCALDRVLKNTA-----WAHSLDAODSLKIWFNNKWHSMVAFVNRASN 1446
Db 795 -----NVLKSPISAFGTAPPMWRKIAVRRAAQVYFNKNGHSMPTYLNSLN 842

Qy 1447 AILRAHLPP-----GPARHAHSITTLNHLNLTKEQLSEALMASSVDVLVLSICVVPAMSFV 1503
Db 843 AILRANLPSKGNPA--AYGIVTNHPNPKTSASLSLOYLL-QGTDVVAIFVIVAMSFV 899

Qy 1504 PASFTVLVIBERVTRAKHLQMLGGLSPPLYLWNLGNLWDMCNVLPACTIVLIFIAFOORA 1563
Db 900 PASFVFLVAEKSTKAKHLQFVSGCNPIIYWLANYVMDMLNLYLVPATCCVILFVFDLPA 959
```

```
Qy 1564 YVAPANLPALILLLLYGSITPLMYPASFPFSPVSTAYVVLTCINLFIGINGSMTATFVL 1623
Db 960 YTSPTNFPVAVLSLFLYGSITPIMYPASFWEVSPASSAVFLVINLFIGITATVATFLL 1019

Qy 1624 ELFS-DQKLEQVSRILKQVFLIFPHFCIGRGLIDVNRNQAMADAFERLGD-RQFQSPLRW 1681
Db 1020 QLFEHDKDLKVNSYLSKCLIFIPNYNLGHGLMAYNEYINYYAKIGQFQDKMSPFWE 1079

Qy 1682 EVVGKNTLAMYIOGFLFTLLLOH-----RSQLLPQPRVRSPLPLLGEDEDEVARREVR 1737
Db 1080 DIVTRGLVAMAVEGVVGFLLTIMCQYNFLRRPQMP---VSTKPV--EDVDVASERQV 1134

Qy 1738 VOGATQGVDLVLRNLTKVYRGO---RMPAVDRCLCIGIPGECFGLLVGNVAGKTSIFRMV 1794
Db 1135 LEGDADNDVKNLTKVYKSKIGRILLAVDRCLGVRFGEFGLLVGNVAGKTSIFXWL 1194

Qy 1795 TGDTLASGEAVLAGHSVARBPSAAHLSGVYCPQSDAIFELITGREHLELLARLRGVPEA 1854
Db 1195 TGDESTTGGEAFVNGHSLKELLQVQSIGYCPQCDALFDELTAREHLQLYTRLRIGISWK 1254

Qy 1855 QVAOTAGSGLARLGLSWYADRPAGTYSGNKKLATALVCDPAVVLDEPTTGMDPSA 1914
Db 1255 DEARVVKWALEKLELTKYADKPAGTYSGNKKRKLSTALIGYPAFIFLDEPTTGMDPKA 1314

Qy 1915 RRLWNSLLAVVREGRSVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPHLLKGRFAAGH 1974
Db 1315 RRLWNLILDLIKGRSVLVLSHSMEECEALCTRLAIMVNGRLRCLGSLQHLKNRFGDY 1374

Qy 1975 TLTRVPAARS-QPAAAFVAAEFPGSELRHAHGRRLQLPFGRCALARVEGLAVHGA 2033
Db 1375 MITVTKSSQSVKDVVRFNRFNPEAMLKERHHTKVQQL-KSEHISLAQVFSKMEQVSG 1433

Qy 2034 EHGVEDFSVQTMLEEVFLYSKDGCKDEDTBEOKEAGVGVDPAPGLQHPKRVSVQPLDDP 2093
Db 1434 VLGIEDYSVSTOTLDNVFVFAK---KQSDNLEQOE---TEPPSALOSPLGCLLSLLRP 1486

Qy 2094 STAETVL 2100
Db 1487 RSAPTEL 1493
```

RESULT 3

```
BS4774
ATP binding cassette transporter ABC2 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 16-Aug-2004
C:Accession: BS4774
R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A:Reference number: A54774; MUID:94375008; PMID:8088782
A:Accession: BS4774
A:Molecule type: mRNA
A:Residues: 1-1472 <LUC>
A:Cross-references: GB:X75927; NID:g495258; PIDN:CAA53531.1; PID:g495259
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:44-234/Domain: ATP-binding cassette homology <ABC1>
F:161-68/Region: nucleotide-binding motif A (P-loop)
F:1108-1300/Domain: ATP-binding cassette homology <ABC2>
F:1126-1133/Region: nucleotide-binding motif A (P-loop)
```

```
Query Match      24.7%; Score 2694.5; DB 2; Length 1472;
Best Local Similarity 40.8%; Pred. No. 1.3e-160;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 750 VEEAPPGLSPGVSVRSLEKRPFGSPQALRGLSLDFYQGHITAFIAGHNGAGKTTLSILS 809
Db 16 MEEEPHTLPLVVCVDKLTQVYKNDKLANKLSLNLYENQVVSFLGHNAGKTTTNSILT 75

Qy 810 GLFPPSGSAFILGHDRVSSMAAIRPHLGVCPQVNVLFDMLTVDHVVYFGRKLGSAAV 869
Db 16 MEEEPHTLPLVVCVDKLTQVYKNDKLANKLSLNLYENQVVSFLGHNAGKTTTNSILT 75
```


QY 261 LMAQVNRTPBELTLRDVREVMELGPRIETFMNDSSNVAMLRLLQMODEGRORPPGG 320
D 9 LLLKKNYTLQKRKVLVTVLELFL---PLLF---SGILLRLKIQSENVFNATYPG- 59
QY 321 RDHMEALRSFLDPGSGGYSWQ---DAHADVGHVLTGLQRTVTECLSLDKLEAPSEALV 376
D 60 -QSIQELPLFFTPPPGDTWEAYIPSHSDAAKTVTETVRRALVINM-RVRGFPSEKDFE 117
QY 377 SRALQLLAEBHRWAGVVLGPDSDPTEHPDPLGPGHVRIRKIMWDIDVTR----- 429
D 118 DYIRDNCSSSVLAIVFEHPNHS---KEPLPLAVKYHLRFSYTRNTYMTQTSFFLK 174
QY 430 -----TNKIRDFWDPGPAADPLT-----DLRYVMGPGFYVLQDLVERAAVRVLSGAN--- 476
D 175 ETEGWHHTSLPLFPNPGP-RETPSDGEGPGVIREGFLAVQHAVDRAIMEYHADAATRO 233
QY 477 --PRAGLYLQMPYCYVDVDFVLVLSRLPLFLTLAWIYSVTLTVKAVRKEKTRLDRT 534
D 234 LFQRLVTVIKRPFPFPFIADPFLVAIQQLPLLLLSFTYTTALTITARAVVQEKRLKEY 293
QY 535 MEAMGLSRVLWGLFSLCLGPFLLSAALLVLVLKLG-----DILPYSHPGVVVFLFLAAF 589
D 294 MRMGSLWLSWNSANFLFLFLLIAASFWTLFCVKVKNVAVLSRSDPSLVLAFLCF 353
QY 590 AVATVYQSFLLSAFTSRANLAAACGLAYFSLYPLVLCVAVWRDLPGAGRVAASLLSPV 649
D 354 AISTISFSPMVSTFSKANMAAAGFLYFTYIYFVFAVRYNMTLSQKLCSCLLSNV 413
QY 650 ARGFCESLALLEBOGQWQWNVGTRPTA-DVFSLAQVSGLLLLDAILYGLATVYLRV 708
D 414 AMAMGAQLIKFPAKGMGTQMRDLSLVNVDVDFCGQVLGMLLLDSVLYGLVTVYMEAV 473
QY 709 CPGYGIPEPWNFPFRRSVMCGPRPKSPACPTPLDPKVLV-----EAPPLGSPGVSV 763
D 474 FPGQGVQVQWTFPTMPSTWCG-KPRVAGKEEDSDPEKURNEYFEABEDLVAGIKI 532
QY 764 RSLEK--RPPGSPQALRGLSLDFYQGHITAFILGHNGAGKTTLSILSLGFLPPSPGSAFI 821
D 533 KHLKSVFRVGNKDRAVRDLNLNLYEGQITVLGHNGAGKTTLSMLTGLFPPTSGRAVI 592
QY 822 LGHDVRSWMAAIRPHLGVCPQVNVFLDMLTVDHWFVGRGLKGLSAAVVPQDRLQDV 881
D 593 SGYEISQDMVQIRKSLGLCPQDHIDFLDNLTVAHLVYFAQLKGLSRQKCPBEVQMLHII 652
QY 882 GLVSKOSVQTRHLSGMORKLSVATFVCGSOVWILDEPTAGVDPASRGIEWELLKYRE 941
D 653 GLEDKWSRSRFLSGMRKKSIGIALTAGSKVLILDEPTSGMDAISRAIWDLLQOKS 712
QY 942 GRTLILSTHLLDEALLGDRVAVVAGGRCLCCGSPFLRRHLGSGYILTVKARLPLTN 1001
D 713 DRTIVLTTHFMDADLLGDRIAMAKGELQCCGSSFLKQKYAGYHMTLVKE----- 765
QY 1002 EKADTMEGSDVTRQEKNGSGSRVGTPTQLLALVQHVWPGARLVEELPHELVVLVPTG 1061
D 766 -----PHCNPEDISQLVHHVFPNATLSSAGAELSFILPRES 802
QY 1062 AHGDSFATLFRDLTLRLAELRTGYGSDTSLEELFLKV----- 1100
D 803 TH--RFEGLFALEKKQKELGASFGASITTMEEVFLRVGKLVDSMDIQALQALVQV 860
QY 1101 VEECAADTMDGSCQQLCTGIAGLDVTLRLKMPPEQTALENGEPAGSAPTDGSGPD 1160
D 861 HERRASDWNVDNLG-----AMDPDGI GALIEER-----T 893
QY 1161 AVGRVQWALTRQIQALLKRLFLARRRGLFPAQIVLPALFVGLVALVFSLIVPPGHY 1220
D 894 AVKLTGLALHQQFWAMPFLKAAVSWREWKVAAQVLPVLCVTALLAINYSSELFD 953
QY 1221 PALRLSPMYAQVFFSDEADPGDGRARLLLEALLQEALEPPVQHSRFSAPFVPAE 1280
D 954 PMLRLTLGEYGRVTVFVS-----PETSOLQOOL-----SEH----- 985

QY 1281 VAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTGSVEVVQNLTRN 1340
D 986 -----LKDALQAEQ--QEPREVLG-----D 1003
QY 1341 LSDFLVKYTPRVLVROGLTKKVVNVRVYGGFSLGGRDPLGSGQELGRSVEELWALLSPL 1400
D 1004 LEEFLI-----FRASVEGGGFN----- 1020
QY 1401 PGGALDRVLKNLTAWAHSDDAOSDKIWNNKGMWSMAFVNVRASNAIRAHLPDGPARG 1460
D 1021 -----ERCL--VAASFRDVGERTVNNALFNNOQAYHSPTALAVVDNLLFK--LLCGP--H 1069
QY 1461 AHSITTLNHP-----LNLTKQELSEALMASSVDLVISCVVPFAMSFVPASFTVLIER 1515
D 1070 A-SIVVSNFPQPSRALQAQKQFNEG---RKGFIDALNL--LFAMAFLASTFSILAVSER 1123
QY 1516 VTRAKHLQIMGLSLPTLWLNFLMDMNYLVPACIVLVILFLAFOBARVAVANLPAALL 1575
D 1124 AVQAKHVQFVSGVHVASFMLSALLWDLISFLPSLLLVVFKAFDVRAFTRDGHMADTLL 1183
QY 1576 LLLLYGWSITPLMYPASFPFSPSTAYVVLTCINLPIGINGSMATFVLEL---FSDQKLO 1632
D 1184 LLLLYGMAIIPMLYMLWNFFLGAATAYTRLTIFNLSGI----ATFLMVTIMRIPAVKLE 1239
QY 1633 EVSRILKQVFLIPHFCLGRGLIDMVN-----QAMADAFERLGRQFQSP-LRWE 1682
D 1240 ELSKTLDDHVLVPLNHLGMAVSSPVENYETRYCTSSSEVAHYCKYKNIQOENFYAMS 1299
QY 1683 V--VCKNLLAMVIOG-----PLFLFTLLLOHRSOLLPOPRVR-----SLPLGLED 1727
D 1300 APGVGRFVASMMAASCAYLILFLIETNLLQRLGILCALRRRTLTLYTRMPVL-PED 1358
QY 1728 EDVARERERVQCATOG---DVLVRLNLTKVYRGQMP--AVDRLCIGIPPGCEGCLLV 1782
D 1359 QDVADERTRILAPSDPSLHTPLIIKELSKYVE-QRVPLLAVDRLSLAVQKCEGCLLV 1417
QY 1783 NGAKTSTFRMVTDTLASRGEAVLAGHSVAREPSAAHLSMOCQSDAIFELLTGREHL 1842
D 1418 NGAKTSTFRMVTDTLASRGEAVLAGHSVAREPSAAHLSMOCQSDAIFELLTGREHL 1477
QY 1843 ELLARLGVPEAOVATAGSLARLGLSWADRPAGTYSGGNKRKLATATLALVGDPAVVF 1902
D 1478 VNYARLGIPERHIGACVNTLRLGLLPHANKLVRTYSGGNKRKLSTGIALIGEPAVIF 1537
QY 1903 LDEPTTGMDPSARRRLLWNSLLAVVREGSVMLTSHSMECEBALCSRALMVNGRPFCLGS 1962
D 1538 LDEPTTGMDPSARRRLLWNSLLAVVREGSVMLTSHSMECEBALCSRALMVNGRPFCLGS 1597
QY 1963 POHLKGRFAAGHTLRLVPAARSQPA-----AAFVAAEPFSGSELREAHGRLRFPQLPPGGR 2018
D 1598 POHLKSGSGVSLRAKQVSEQQEALBEFKAFVDLTTPFGSVLEDEHQGMVHYHL-PGRD 1656
QY 2019 CALARVRELAVHGAHEGVEDPSVQTMLEEVFLYFS 2055
D 1657 LSWAKVFGILEKAKYGVDDYSVQSISLEQVFLSFA 1693

RESULT 5

A59188

ATP-binding cassette transporter ABC3 - human

C:Species: Homo sapiens (man)

C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004

C:Accession: A59188

R:Connors, T.D.; van Raay, T.J.; Petry, L.R.; Klingler, K.W.; Landes, G.M.; Burn, T.C.

Genomics 39, 231-234, 1997

A:Title: The cloning of a human ABC gene (ABC3) mapping to chromosome 16p13.3.

A:Reference number: A59188; MUID:97179225; PMID:9027511

A:Accession: A59188

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1704 <CON>

A:Cross-references: UNIPROT:Q99758; GB:U78735; NID:gl699037; PIDN:AAC50967.1; PID:gl699

C:Genetics:

A;Description: The sequence of *C. elegans* cosmid Y39D8C.

A;Reference number: 221408

A; Accession: T33783

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-1802 <BEC>

A;Cross-references: UNIPROT:Q9TXV8; EMBL:AF101313; PIDN:AAC69223.1; GSPDB:GN00023; CESP:

A; Experimental source: strain Bristol N2: clone Y339p8C

C;Genetics:

A;Gene: CESP:Y39D8C.1

A;Gelle: CESP:139D8
A:Map position: 5

A:Map position: 5
A:Introns: 45/3: 114/1: 195/1: 230/3: 543/3: 794/1: 849/1: 1036/2: 1099/1: 1132/3: 1165/1

C; Superfamily: ATP-binding cassette homology
A; Intron8: 45/3; 114/1; 195/1; 230/3; 543/3;

Query Match	16.8%	Score 1829;	DB 2;	Length 1802;
Best Local Similarity	28.4%;	Pred. No. 4e-106;		
Matches 539;	Conservative 312;	Mismatches 667;	Indels 382;	Gaps 55;

QY	1102	BECAADTDMEDGSCGQHLCCTGIAGLDVTILKMPQPOTALENGEPAGSAPETDQSGSPD-	1166
DB	907	HTADERYNYEHG-----IEN--DISEMIEKDDPILOQLD	937
QY	1161	AVGRVOCWALTROOLQALLKRFLARRRRRGLFAOIVLPALFVGLALVFSLIVPPFGH	1219
DB	938	RAQVRVTGFTLQMHAKAMFYKRAIFPFRKWTQELPOLVFPVAVILVLMVTSQVLP-	993
QY	1220	YPALRLSPTWYGAQVSFSEDAQDQGRARLLEALLQEAGLEBPVQVHSHRFSAPVPA	1279
DB	994	-----SVKE--	997
QY	1280	EVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGSVEVVQNLTG-	1338
DB	998	-----QDPQITISLAPFSDTKKAG-----HLVSDSGNYVTLTGG	1031
QY	1339	RNLSDFLVKTPRL-VRQGL-----KTKKWNVRYGGFSLGGRDGLPSQBELGRSVEE	1392
DB	1032	QNLSSMWQGVTVQLGVTVQVDTISNVEKEFIMQTN--ANGERTFGLHYA-----	1079
QY	1393	LWALLSPGGALDRVLKNIJAWAHSLDAQDSLKIWNPNKNGHWSVAFVNRASNAILRAH	1452
DB	1080	-----LGFVP-----SMENFSTVSVPSLKTFPNFGLVTPALAITFTDSMILLSQ-	1123
QY	1453	LPQCPARHAHSITTLNHL-----NLTKQLESEA--LMASSVDVLVSI CVVPAMSVFP	1504
DB	1124	-----KOKKQVSTAVNHPPLPPTODTLKNTNRSDGAFLIAYGLIVSFAVCV-	1171
QY	1505	ASFTLVLIBERVTRAKHLQMLGSLPTLYWGLNFLWDMCNLYLPACIVLILFLAFQORAY	1564
DB	1172	AGVSQFLITERKKSKHMLLSGIRPWFMTAFIWDAAWFVIRILCFDAIFIFINITAY	1231
QY	1565	VAPANLPALLLL-LLLYGNSITPLMYPASFFSVPSTVYVVLTCINLFIGINGSMAFVL	1623
DB	1232	THDPGVMLILTSLFLYGTALPFTYWFQFFESAPKGFMMVTWYHILTCMIGSIAVPII	1291
QY	1624	E-----LPSDKLOEVSRI-----LKQVFLIPHF-----CLG	1651
DB	1292	SOFTSSLDAGYLMIIIPAWLPFTYNIQIATVTFQENVRJACKKLOCTIIMFKAVTACCG	1351
QY	1652	RLGLDMVRNQAMADAPERL-----GDRQFOSPLRWVVGKLLAMVI-----QGFLPL	1699
DB	1352	-----TASERLVONVLFVGNR-----KGILVYIFLAVQGFYW	1386
QY	1700	LFITLLLOHR--SOLLPOPRVRSI--PLLGBEEDVARERE-----RVVQATQGD	1745
DB	1387	IWFMRENDOFTKLIFALIRCKADNPITDITDKYDERVEDSDVIAEKSVVQRLANN	1446
QY	1746	--VLVLNLTKVYRGQMPAVDRCLGIPGCEGFLGVNGACKTSTFRMVTGDTLASRG	1803
DB	1447	KTALVSNLNVKMY--GNFNKAVGNVFNHNSKOCFGLLVNGAGKTFQMLTGENSISG	1504
QY	1804	EAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLITGREHLELLARLGRVPEAQVATAGSG	1863
DB	1505	DAYUNGWSVQNNREAGANTGYCPQYDAIIKENSGEETLYMPARIIGIPEKDIPKKNV	1564
QY	1864	LARLGSWYADRPAGYVSGNKKKLATALVGDPAVFLDEPTTGMDPSARFLWNSLL	1923
DB	1565	IHAIGIMWASROIKTYVSGNKRRLSIGIAIVGLPDVLLIDPTSGVDPKARRIINILN	1624
QY	1924	AVVREGSVMLTSHSMECEALCRLAIWNGVFRCLGSPQHLKGRFAGHITLTLAVPAA	1983
DB	1625	RLNDGLTALVLTSHSMDCEALCTELAIWYGFRCYGCQHIKSRYGSGYTLILILKNR	1684
QY	1984	R-SQAAAAFVAABFPGSSELRHAGGRLRFOLPGGRCALARVFGELAVHGAEBGVDFSV	2042
DB	1685	NDAEKTSTIKQTFRGSVIKEHVLIQNPDI PRDGD-SWSRFEKLETVSTSLWDDYSL	1743
QY	2043	SQWLBEVFLYFSKQDKOEDTEBEQKAGVGVDPAFGLQH	2082
DB	1744	SQTTLEQVTFEFRDAGVSDSEDFDGSASSTGANSRKGQN	1783

QY 1833 FELLTGREHLELLARLRGVPEAQAQTAGSLARGLSWYADRPAGTYSGGNKRKLATL 1892
Db 1471 FEYLTVKEHLELYARIKGVVDHRIDNVVTEKLVFEDLLKHSHKSPSTLGGNKRKLSVAI 1530
QY 1893 ALVGDPVAVFDEPTGMDPSARRFLWNSLLAV-VREGH-SYMLTSHSMBECCALCSRLA 1950
Db 1531 AMIGOPPIVILDEPSTGMDPVAKRPMWMDVISRLSTRSGKTAVILTTHSMNEAQAALCTRIG 1590
QY 1951 IMVNGFRCLGSPQHLKGRFAAGHTLTLRVP-----AA 1983
Db 1591 INVGGRLRCIGSPQHLKTRY--GNHLELVFPYNGVKPNEVSNVELENFCQIIQQWLFNV 1648
QY 1984 RQCPAA-----1989
Db 1649 PTQPRSLLDLEVCIGVSDSITPTDASASEISLSPMWQRIAKFLGNEQVSTLVPPLPE 1708
QY 1990 -----AFVAAEPGSELREAHGRLRP 2011
Db 1709 EDVRFDDQLSEQLFRDGGIPLPIFAEWLTKFKSALDSFIQSSPFGATFKSCNGLSIKY 1768
QY 2012 QLPPG-GRCALARVFCGLAVHGAEGHVEDFSVSQTMLEEVFLYFS 2055
Db 1769 QLPPGEGGLSLADAFGLHLENNRGLIAEYSISQSTLETIFNHFA 1813

RESULT 8
C88925
protein F33811.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C88925
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C88925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1317 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC17542.1; PID:G3158495; GSPDB:GN00023; CESP:F33811.
C:Genetics:
A:Gene: F33811.4
A:Map position: 5

Query Match 14.7%; Score 1603; DB 2; Length 1317;
Best Local Similarity 28.3%; Pred. No. 3.9e-92;
Matches 447; Conservative 246; Mismatches 511; Indels 378; Gaps 41;

QY 593 TVTQSFLLSAFFSRANLAAACGGLAYFSLYLIPYLCVAVWRDLRPAAGRVAASLLSPVAFG 652
Db 3 SATSNTLISVFWM-----LLYF--WTAFFSSIDQTPYPLGYRLINCINPDIALN 51
QY 653 FGCESLALLEBQEGAQHNHVTREPTAD---VFLSAQVSGLLLDAAALYGLATWYLEAVC 709
Db 52 YGLQLLAAYEQADGLKAGELFTPPSPDNLTTFGHALLA--LIVDGIIMILTWYIEAVI 109
QY 710 PQQYGIPE-PNPFPRRYSWCGPRPKSPACPPLD-----PKVLVEAPPLSPG 760
Db 110 PGGEVPOKPFVFL-PSYWF-----PNSGSKTVSDSQFQIQVADHVKLEKEPTDLIPT 164
QY 761 VVSLEKRFPGS-----POPALRGLSLDFYGHITAFILGHNGAGKTTLS 806
Db 165 INVNLTKYTGTSFFKFLFDCKFGKSGEKAIVSNLKNYPCQCTVLLIGHNGAGKSTTFS 224
QY 807 ILSGLFPSPSGSAFTLGHVRSMAAIRPHLGVCPCQYNVLFDMLTVDHVMFYGRKGLS 866
Db 225 MLTGWASPSGSAVYNDFTSLPKIRRENGLCPCQYNLTGFWMTMEHLEFFAKLK---281
QY 867 AAVGVPEODR-LLQDVGLVUSQSVOTRHLSGMQRKLSVAIAFVGGSQVWILDEPTAGVD 925
Db 282 ERTWDPEEARILARLIDFKADEFMAGALSGGQKRKLSLAIALIGSGSVVMDLDEPTSGMD 341

QY 926 PASRRGIWELLKRYREGRTLILSTHLLDEALLGDRVAVVAGRLCCGSPFLRRHLGS 985
Db 342 POARHETWLIQREKERTILUTTHFMEADLLGRIAMAGQLECCGSPFLKQOYGD 401
QY 986 GYYTLVKARLPLTTNEKADTMEGSDVTRQEKQSGSQSRVGTTPOLLALVQHVPGARL 1045
Db 402 GYHLTIV-----YDTTSTPDVSKTD-----IIREYIPEAHV 433
QY 1046 VEELPHELVLVLPYTGADGSPATLPRELDTLAEURLTGYIGSDTSLEEILFKVVEBCA 1105
Db 434 FSYIQEATYLLSAT--HRPIFPKFLKELEDHQTGCGTSFGVSTTMEVEFLKVGHGTAD 491
QY 1106 ADTDMEDSCGHLCTGIAGLDVTLRLKMPPTALENGEPAGSPAPETDQSGPD--AVG 1163
Db 492 BRYNYEHG-----IEN--DISEMIEKDDPILQDLRAQV 522
QY 1164 RVQGWALTROQLQALLKRLARRSRRLFAQIVLPALFVGLALVFLSILVPPFGHPAL 1223
Db 523 RVTGFTLQWQHAKAMFYKRAIFFRKTQFLPOLVFPVAYLVLMVFTSQVLP-----574
QY 1224 RLSPTMYGNOVSFFSEDAPGDPGRARILLEALLQEALEPPVQSHSRPSAPEVPAEVAK 1283
Db 575 -----SVKE-----578
QY 1284 VLAGNWTMPSPSPACQCSQCGARLLPDCPAAAGPPPPQAVTGVGGEVQNLTG-RNLS 1342
Db 579 -----QDQFOTISLAPFSDTKKAG-----HLVDSGNVTVLGGSQNLS 616
QY 1343 DFLVKTYPRLVROGLTKKWNVRYGFSGLGRDPLGSPQELGSRVSEELWALLSPLPG 1402
Db 617 SMV-----QGTVTQLGVQT-----TVVDITSNVEK-----641
QY 1403 GALDRVLKULTAWAHSLLDAQDSLKIWFNNKWHSMVAFVNRASNAIILRAHLPPGPARHAH 1462
Db 642 -----FIMDQTNAM-----GS 652
QY 1463 SITTLNHLNLTKEQLSEALMASSVDVLSTCVVFMASFVPASFTLVLIBERTVTRAKHL 1522
Db 653 RTFGLHYALGFVSPMFNFSTVSFSLKISPAVCV-----AGYQFOLITERKKSRHM 704
QY 1523 QLMGSLSTLYWLNFLNMDNCLVLPACIVVLIFLAQOQRAYVAPANLALLL-LLLYG 1581
Db 705 QLLSGIRPMPFWLTAFIWDAAFWIRILCFDAIFYFNITAYTHDFGVMILITLSFLLYG 764
QY 1582 WSITPLMYPASFFSPSTAYVVLTCINLFTGINGSMATFVLE-----1624
Db 765 WTALPFTTYWQFPFESAPKGFMMVTMYHILTMIGSIAVPIISQTSLSLDAGYLWSIIFAW 824
QY 1625 LPSQDKLQEVSKI-----LKOVELIIPHF-----CLGRGLIDMVRNQAMADAPER 1669
Db 825 LFPTYNISQIATVTFQENVTIRACKLDCTIPMFKAVTACCG-----TASER 871
QY 1670 L-----GDRQFSPLRWEVWGKNLAMI-----QGFLFLFTLLLQHR--SOLLPOP 1715
Db 872 LYVDNVLFPVGNR-----KGLVYVIFLAVQGFYIWMVFMRENDQFTKLPAI 919
QY 1716 RVRSJ--PLGSEEDVDVARE-----RVVQATQGD--VLVRLNLTKVTRGORM 1761
Db 920 RCRKADNPIMDITDVKYDERDVEDSDVIAEKSVVYQRLANNKKTALVSNLWKVY--GNF 977
QY 1762 PAVDRLCIGIPPGCFGLGVNGAGKSTTFRMVGTDLASRGEAVLAGHSVAREPSAHL 1821
Db 978 NAVGVNFHVNKSCFGLLVNGAGKSTTFQMLTGENSISISSGDAYVNGSVNNWREAGA 1037
QY 1822 SMGYCPOSDAIFELLTGREHLELLARLARGVPEAQAQTAGSLARGLSWYADRPAGTY 1881
Db 1038 NTGYCPOYDAIKEMSGEETLYMFARIRGIPEKIDIPKKNVAVIHAIGIMVARSQIKTYS 1097
QY 1882 GGNKRKLATALLVGDPAVFLDPTTGMDSARRFLWNSLLAVVREGRSVWLTSHSMBE 1941
Db 1098 CGNKRRLSLGIAIVGLPDVLLDPTSGVDPKARRIWIINILNRLDLGTLVLTSHSMD 1157

Db 1443 -GRFVAVNELCLAVDQKCEFCGLLVNGVAGKTTTNTLLTQSGFASGEAMIGRDVTE--- 1498
Qy 1818 AAHL5MGVCPQSDAIFELLTTOREHLELLARLURGVPEAQVATAGSLARLGLSWYDRPA 1877
Db 1499 --LISIGYCPQFDALMLDLTGRSLEILAAQHGHPENYKA--KABELILECVGMIADKLV 1554
Qy 1878 GTYSGNKRKLATALALVGDPAVFLDEPTTGMDPSARRFLWNSILLAVVREGS-VMLTS 1936
Db 1555 RYSGGQKRIKISVGALLAPTQMIILDEPTAGIDPKARREVWELLMLWCREHNSALMLTS 1614
Qy 1937 HSMDECEALCSRLAIMVNGRCFLGSPQHLKGRFAAGHTTLRV--PAARSQPAFAVAA 1994
Db 1615 HSMDECEALCSRIAVNLSGSLIAIGSSQELKSLGNNYTWLTSLEYENQDM-VVOLVQT 1673
Qy 1995 EFGSELRHAGGR---LRFOLPGRCALARVFGELAVHGAHGVDFSVQTMLEBEVF 2051
Db 1674 RLPNSVLKTTSTNKTLLKWIPIKEDCWSAKPEMVQALAKDLGVKDFILAQSLSBETF 1733
Qy 2052 LYFSKDOQKDED 2063
Db 1734 LRLA---GLDED 1742

RESULT 11
S60124
transport protein homolog C48B4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 13-Jan-1996 #sequence_revision 12-Apr-1996 #text_change 16-Aug-2004
C:Accession: S60124; S40724; S40725
R:Kershaw, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: S60124
A:Accession: S60124
A:Molecule type: DNA
A:Residues: 1-1767 <KER>
A:Cross-references: EMBL:Z29117; NID:9439247; PID:g1066912
C:Genetics:
A:Map position: III
A:Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein
F:628-818/Domain: ATP-binding cassette homology <ABC1>
F:645-652/Region: nucleotide-binding motif A (P-loop)
F:764-769/Region: nucleotide-binding motif B
F:1457-1642/Domain: ATP-binding cassette homology <ABC2>
F:1474-1481/Region: nucleotide-binding motif A (P-loop)
F:1586-1591/Region: nucleotide-binding motif B

Query Match 12.9%; Score 1407; DB 2; Length 1767;
Best Local Similarity 25.9%; Pred. No. 1.2e-79;
Matches 485; Conservative 298; Mismatches 665; Indels 424; Gaps 59;

Qy 337 GYSQDAAH--DVGHVGLTGLRVTECLSLD---KLEAAPSEAAVLSR-ALQLLAE----- 385
Db 159 GYTSKDAAAKTVDDLMKKFAERQSAKLLSVKNSSEEEQLTLVRLNDLPMLNETFCAL 218
Qy 386 HRFWAGVVF-----LG---PESSDPTSHPTDPLGPHVIRKIRMDIDVVT 428
Db 219 NSYAAGVVDFVDVNTKKNLYRIILGKTPEETHLTETSYNYPGSSGR----- 267
Qy 429 RTNKIRDFDPGPAADPLDRLRVYVWGFFVQLDVERAAVRLSGANPRAGLYQOMPY 488
Db 268 -----YRIFSSPP-----YWTSALFTQFAIESFLSSVQSGSAPDLPITLRGPE 313
Qy 489 PCYVDDVFLVLSRLPLFLTLAMIYSVTLTVKAVRE---KETRLRDTMRAMGLSRAVL 545
Db 314 PRYKTSVSAFID-----FPPPIAFVTFINVIHITREIAENHAKVPLTAMGLSTFMF 368
Qy 546 WLGNFLSCLGPFLLSALLVLVLKGLDILPYKSHPGCVFLFLAAFAVATVQTOSLLSAPFS 605
Db 369 YAAHVWMAFLKFFV--IFLCSIIPLTFVMEFVSPPAALITVITMYGLGAVIFGAFAFFN 426
Qy 606 RANLAAACGGLAYFSLYLPYVLCVAVMRDL-----PAGGRVAASLLSPV-----AFGFG 654

Db 427 NTNSAIK-----AILVANGAMIGISKLRPELDQISSCFLYGLNINGAFALA 473
Qy 655 CESLA-----LLEQEGGAQWNVGTRPTADVFSIAQVS-GULLLD-----AALYG 699
Db 474 VEAISDYRRERELNTNMFNDSLSHPSLWALVMIMVIDILMWSIGALVVDHRTSADFS 533
Qy 700 LATWY-----LEAVCPQGYGIPPEWNPFRFESYWCWGPSPKSPAPCPTPLDPKV 748
Db 534 LRTLFDPEAPEDDENQDGTVAQNTRIQEQVRNRVRSD-MEMPMASSTSLNPNADSDS 592
Qy 749 LVE-----BAPCLSPGVSVSLKRPFGSPOLARGLSLDFYQGHITAFGLHNGAGKT 802
Db 593 LLEGSTADGARDTARADIIVRNVLKIVSTTGERAVDGLSLRAVRGQCSILLHNGAGKS 652
Qy 803 TTLSILSGLPPPGGSAFIIIGHDVRRSMAAIRPHLGVCPQNVNLFDMLTVDHVMVFGRL 862
Db 653 TTFSSIAGIIRPTNGRITICGYDVGNPEPGETRRRHIGMCPQYNPLVDQLTVSEHLKLVGL 712
Qy 863 KGLSAAVVGPEQDRLLQDGLVSKQSVOTRHLSCGMQRKLSVAIAFVGGSQVVLDEPTA 922
Db 713 KGAREKDFQDMKRLSLSDVKLDFKENEKAVNLGGMKRKLVCVCMALIGDSEWLLDEPTA 772
Qy 923 GVDPASRRGIWELLKYREGRTILILSTHLDLAEELGDRVAVVAGGRLCCGSPFLFRRH 982
Db 773 GMPGARQDYKLVVEREKANRTILITTHYDDEARLGDWFIIMSHGKLVASGTNQYLKQK 832
Qy 983 LSGGYILTLVKARLPLTNEKADTMEGSVDTTRQEKNGSQGSRVGTGPQLLALV-----Q 1037
Db 833 FGTGYLLTV-----LDHNGDKRK-----MAVILTDVCT 861
Qy 1038 HWVPGARLVEHPHELVLVLPYTCAGHDSFATILFRELD-----TR 1077
Db 862 HYVKEAERGEMHQOQIBIILP--EARKKEFVPLFQALEAIQDRNYSRVFNDNMNTLSQ 919
Qy 1078 LAELRLTGYGIDSTSLEIEIFLVKVEECAADTMEDGSGQHLCTGIAGLDVTLRLKMPQ 1137
Db 920 LATLEMRSGLSLNTLEQVITI-----GDKDKAIASRQ-NSRISHNSR 963
Qy 1138 ETALENGEPAG-----SAPETD-----QGGSPDAGRVQGWALTRQOLQALLKRFL 1184
Db 964 NASEPSLKAGYDQTQSTKSDSYKQLMDSQARGPEK-----SGVAKKVAQFISIMRKKFL 1019
Qy 1185 LARRSRGLFAQIVLPALFVGLALVFSLIIVPPGHYPALARLSPTMYGAQVFFSEDPAGD 1244
Db 1020 YSRNWAQLFTQVLIPILILGLV-----GSLTTLKSNNT-----D 1054
Qy 1245 PGRARILLEALLQEALEPVPQHSRFSAPAEVPAEVAKVL-ASG-----NWTPESSPA 1298
Db 1055 QPRS-----LTPSGIIPSKVVMWRFPENGITPEEAAEFKILRKSQGFVLNNTKNP--- 1105
Qy 1299 CQCSQPGARLLPDCPAAAGGPPPPQAVTSGGEVVQNLTRNLSDFLVKTYPRILVROGLK 1358
Db 1106 -----LPNITKSLIGEMPATI----- 1122
Qy 1359 TKXWNEVRYGGSLSGORDPGLPSGQELGRSVEELWALLSPLPGALDRVLKNTAWAHS 1418
Db 1123 -----GMTMN----- 1127
Qy 1419 LDAQDSLKIWFNNKGWH---SMVAFVNRAS-NATLRAHLPPGPARHAHSITTLNHPNLIT 1474
Db 1128 ---SDNLLEALFNRYHYHVLTLISMNRLRTGTVDAEISSGVFLYSKTSN-----SNLL 1180
Qy 1475 KEQLSEALMASSVDVLVSVICVVFAMSFVPASFTVLIERVTRAKHLQLMGGLSPFLYW 1534
Db 1181 PSQI-----IDVLLAPMLILIFAMVTSTFVMEFLIEERTCOFAHQQLFGTGISPTFY 1231
Qy 1535 LGNFLWDMCNVLPACIVLILFAPQORAYAPANLPALLILLLLLYGWSITPLMYPASFF 1594
Db 1232 SASLIYDGIIYSLICLIFLFWFLAF-----HWMYDHLAIVILFWFLYFSSVPFIYAVSFL 1287
Qy 1595 FSVPTAYVVLITCINLPIGINGSMATFVLSEFSDQKLOE-VSRILKQVFL-IFPHFCLGR 1652

1534 WLGFLDMCNVLPACIVLVIFLAFQRAVAVAPANLPALLLLLLLLYCWSTITPLMPASF 1593
1568 YSALIVDGLYSLICLIFLFWLAF-----HWYDHLAIVLFWLFFSVPPIYAVSF 1223
1594 FFSVPSTAYVLTICINIFIGNGSMATFVLFSFQKLOE-VSRILKQVFL-IPFHFLG 1651
1224 LFQSPSKANVLL--LIQVVISGALLAVFLIFWIFNIDWLSKILVNIWFMLFSPVAFG 1281
1652 RGLDMVRNQAMADAFERLGRQFQSPPLRVEVGNLLAMVIOGFL-FLFLTLL----- 1704
1282 SAIL-TINTYGMILPSEEL-----MNWDHCGKNALMGTFGVCSPALFVLQFKVR 1332
1705 -----LQHSQLLPQPRVRSPLLEGEDEDVARERERVQCATQGDVVLNLTQVY 1756
1333 RFLSQWTVRRSSHNVPMMGDLFVC-----ESYSEERVRHVRNSONSALVIKDLTKTF 1388
1757 RQMPAVDRCLGIPGCEFGLLGVNGAGKTSIFRMVGTTLASRGEAVLAGHSVAREP 1816
1389 --GFTAVNELCLAVDQKECFGLGVNGAGKTTTFNLTGQSFASGEMIGGRDVTB-- 1444
1817 SAHLSMGYCPQSDAIFELITGRHELLARLGRVPEAQVAQTAGSGLARLGLSWYADRP 1876
1445 ---LISGYCQFQDALMLDTGRESLEILAQMHGFFENYKA--KAELILECVGMIAHADKL 1499
1877 AGTSGGNKRLATALLVGPVAVFVLDPEPTGMDPSARRLWNSILLAVVREGS-VMLT 1935
1500 VRFYSGGQKRKISVGVALLAPQMIILDEPTAGIDPKARREVWELLWCREHSNSALMLT 1559
1936 SHSMECEALCSRLAMVNGFRCLGSPQHLKGRFAAGHTLTLRV--PAARSQPAAPFA 1993
1560 SHSMECEALCSRLAVLNRSGLIAIGSQELKSLYGNNTWLTSLYEPNQDM-VVQLVQ 1618
1994 AEPFSGELREAHGR---LRQLPPGRCALARVFGELVAGHGVDEFSVQTMLEEV 2050
1619 TRLPNSVLKTTSTNTKLNKQIPEKEDECSAKFEMVQALAKDLGVKDFILAQSSLEET 1678
2051 FLYESKQOGKDED 2063
1679 FLRLA---GLDED 1688

RESULT 13
T00826
hypoetical protein T3266.22 - Arabidopsis thaliana (fragment)
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00826
R/Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A/Description: Arabidopsis thaliana chromosome II BAC T3266 genomic sequence.
A/Reference number: 214163
A/Accession: T00826
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1246 <ROU>
A/Cross-references: UNIPROT:Q088T6; EMBL:AC002510; NID:g2618683; PID:g2618705
A/Experimental source: cultivar Columbia
C/Genetics:
A/Map position: 2
A/Introns: 33/3; 95/2; 113/3; 137/3; 168/3; 361/3; 421/2; 432/3; 493/3; 521/3; 535/3; 56
A/Note: T3266.22

Query Match 12.6%; Score 1376; DB 2; Length 1246;
Best Local Similarity 28.7%; Pred. No. 6.4e-78;
Matches 403; Conservative 195; Mismatches 415; Indels 392; Gaps 42;

881 VGLYKSOVQTRHLGGNOKRLSVIAIFVGSQVVIIDEPAGVDPAARRGINELLILKYR 940
1 VGLSDKINTLVRALSGGKRLKSLGIALIGNSKVIILDEPTSGMDPYSMRLTWOLIKIK 60
941 EGRLLIISTHLLDAEALLGDRVAVVAGRLCCGSPFLRRHLGSGYLLTVKARLPLTT 1000
61 KGRILLTTHSMDAEELGDRIGIMANGSLKCCGSSIFLKHHYGVGTILVKK----- 113

QY 1001 NEKADTMESVDTRQEKNGSQSRVGTTPOLLALVOHWVPGARLVZELPHELVLVLPYT 1060
DB 114 -----TSPTVSAAH-----IVRHIPSATCVSEVGNEISFKLPL- 148
QY 1061 GAHDGSFATLPRELDTRLAEL-----RLTGYGSDTSLEBIFLKV-----VECAADTD 1109
DB 149 -ASLPCPENFREIESCMKNSDSYDYGISYGVITLVEEVFLRVACNLDIEDKQEDIF 207
QY 1110 MEDSCCOHLCTAGLDVTLRLKMPQETALENGEPAG-----SAPE----- 1152
DB 208 VSPDTKSLVIG-----SNQSSMQPKLLASCN-DGAGVITTSVAKAFRLIVAAVWTLI 261
QY 1153 ---TDQSGDPADVGRVCGWALTQOQALLLKRPFLARRRRGLFAQIVIPALFVGLALV 1209
DB 262 GFISIQCCGSIISRSFV-----RHCKALFIKARSACRDKTVAFOFIIIPAVFLFGL 317
QY 1210 FSLIVPFGHYPALRLSPVWGAQVSFSDADPGRRLLLEALLQVLEAGLEEBPVQHSS 1269
DB 318 F-LQKPHPOKSTILT-----TAYFNLLSGKGGG-----PIFPD- 353
QY 1270 HRFSAPVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAG---GPPPPQAV 1326
DB 354 --LSVP-IAKEVAQYI-EGGWI--QPLRNTSYKFPNPKALADAIDAAGPTLGP----- 401
QY 1327 TGSGEVVQNLTGRLNLSDFLVKTYPRLV---RQGLTKKWVNEVRYGFSLGGRDPGLPSG 1383
DB 402 -----TLISMSEFLMSFDSQSYQSSREGLSHSDCNH----- 433
QY 1384 QELGRSVEELWALLSPLPGCALD-RVLKNLTAWAHSLLDAODSLKIWNKNGHWSMAFVN 1442
DB 434 -----PDGSLGYTVLHNGTC-----QHAGPIYIN 457
QY 1443 RASNAILRAHLPPGPARRHAHSITTLNHNPLNLTKEQ-LSEALMASSYVDVLVSIWCVFAMS 1501
DB 458 VMHAAILLR---ATGNKNWTIOTRNHPLPTKTQRIQRHDLDAFSAIIVNI---AFS 509
QY 1502 FVPASFTLVLIBERVTRAKHQLMGSLPTLYLGNFLDMCNVLPACIVLVIFLAFQ 1561
DB 510 FIPASFAPIVKBREVKAHQQLISGVSVLSYMLSTYVWDFISFLPFSTFAILFYAFGL 569
QY 1562 RAYVAPANLPALLLLLLLYCWSTITPLMPASFPFSVPSTA-----YVVLTCINLFING 1616
DB 570 EQFIGIGRFLPTVLMLEYGLAIASSTYCLTFFTEHSMQAOTSSYVLIPSLFVSFS 629
QY 1617 SMATFVLELFSQKLOEVSRI-----LKVQFLIFPHFCLGRG 1653
DB 630 SNVILMVHFFSGLILMVISFVMGLIPATASANSYLKELILFRYALQNFRLSPGCFSDG 689
QY 1654 LIDMVRNQAMADAFERLIG--DROFQSPLRVEVVGKN-----LLAMVIQ 1694
DB 690 LASL-----ALLRQGMKDKSHGVFWNVGTGASICYLGLVLEVCYRYSMLLSFFH 741
QY 1695 G---PLFLFLLTLQLHRSQLLPQPRVRSL-----PLIG-----EDEDVARER 1736
DB 742 GIDTKLSLIYTIIGASRLTELI-YDRVYSTSFTEPLKDKSTGAISTMEDDIDVQERDR 800
QY 1737 VVOGATQGDVVLNLTQVYRQGR---MPADVRLCLGIPPGCEFGLLGVNGAGKTTFR 1792
DB 801 VISGLSDNTMLYLQNLKRVYDGHGPKVAVQSLTFSVQAGCEFGFLGTNGAGKTTLS 860
QY 1793 MVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLQVP 1852
DB 861 MLSGEETPTSGTAFIFGKOIVASPKAIRQIHGICQCPQDALFEYLTVKHELELYARIKGV 920
QY 1853 EAQVAQTAGSLARLGLSVADRPAGTYSGNKKLATALALVGDPAVFLDEPTTCMDP 1912
DB 921 DHRIDNVVTEKLVFEDLLKHSHRPSFTLSGNGKRSKLSVAITAMIDPPVILDEPSTGMDP 980
QY 1913 SARRFLWNSLLAV-VREGR-SVMLTSHSMECEALCSRLAIMVNGFRCLGSPQHLKGRF 1970
DB 981 VAKRFMWVDSIRLSRSGTKAVILTTHSMNEAQLCTRIGIMVGGRLRCIGSPQHLKTRY 1040


```

Db      923  -----FIVKVNK 929
Qy      1351 R-----LVQGLKTKWNEVRYGGSLGRDPLSGQELGRSVEELWALLSPLPGGA 1404
Db      930  KGDSKFYKISQGDKNAAIILMNIITASAMYL--RDPSTVK-----965
Qy      1405 LDRVILKNLTAWNSLDQDSIKWNN-----KGWHSWAFVNPASNAIILRAHLPPGPAR 1459
Db      966  -----LPHVT-----SRVIMNDPKIKYEGLASFFLFEN-----994
Qy      1460 HAHSITTLNHLNLTKEQSEALMASSVDVLVSCVVFAMSFVPASFTIVLIEERVTRA 1519
Db      995  -----IFFLLVLAGIFIQS--TVYLIEEKICKP 1020
Qy      1520 KHLQMLCGLSFTLYWLNFNWDMCNLYLPACIVVLIFLAFQORAYVAPANLPALL--LLL 1577
Db      1021 AHQYLTGLSTIAYWGVVFLWDF-----LLFTFELLYTIGLISFVGLQGHIEIV 1072
Qy      1578 LLYG--WSITPLMYPASFFSPSTAYVLTICNLFIGINGSMATFVLELFSQKQEV 1635
Db      1073 IFYGLLEFYFAPLVLTALINTPTGNFLL---YMFCCIPLWAYSVISELHNFPPFIQKYS 1129
Qy      1636 RILKQVELIF-PHFCLGRGLIDMVR---NOAMADAPERLGRDQFQSPLRWEVVGNLLA 1690
Db      1130 DEIEYGRIFNPSTGFLAGLMKIAALNPKSGLDKHEHLTNLTWYEGIFPE-----LM 1183
Qy      1691 MVIQGPLFLFTLL-----QHRSQLLPQPRVRSPLGLGEDEEDVARER 1734
Db      1184 FLFFGGIFL--TILLGCATLKPFRACFRGTRRSQ--PREKR--YKGIESC KAVKEE 1236
Qy      1735 ERVQOGATQGD-VLVLRNLTKVYRGQMPAVDRCLGIPGCECPGLLVNGAGKTSTFRM 1793
Db      1237 EQLVQVEDRKNETVLVIDGLVKDF--GKFRAVNDLSISVGHEECFGLGANGAGKTTTFDI 1294
Qy      1794 VTGDTLASRGEAVLAGHSVAREPSAAHLSMGCYCPQSDAIFELLTGREHLELLARLGVPE 1853
Db      1295 ITGLTMFTGGSATIDGHITET-----IHIGYCFQPDAMLQQISCRQTLIMAKLOGYP- 1348
Qy      1854 AQVAQTAGSLARLGLSWYADRPAGTYGGNKRKLATALALVGDPAVVFLDEPTGMDPS 1913
Db      1349 -NVKEVVVELVDCVMSDFGYKLVKNGSGQKRIKISVGIALMSRATCIILDEPTAGIDPR 1407
Qy      1914 ARRLWNSLLAVVPEGR-----SYMLTSHSMECEALCSRLAMVNGRFRCLGSPHLLKG 1968
Db      1408 ARREIWD-----IHEMREQA KCSIVLTSMSMECEALCTRIGILRKEMIALGTSQSLKS 1463
Qy      1969 RFAAGHTLTLRVPAARS-QPAAAFVAAEFPGSELREAHGG---RLRFPLPPGRCALARV 2024
Db      1464 QYGNTRYMMTLILNSLEDLESVCVIVSEMPDAVLKTPESSITTSIVWELPKSKDKWSEK 1523
Qy      2025 FGELAVHGAHEGVDFSVSQTMLBEVFLYFSKQGDDEDETEQKEA 2070
Db      1524 YNQVEVLAKKANAKDYMLTQASLEDTFIRLI-----TTEEBEEA 1562

```

Search completed: January 3, 2005, 10:55:35
Job time : 66.9746 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 238.389 Seconds
(without alignments)
5068.549 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHHECHFPNKLPSAGT.....QHPKRVSQFLDDPSTAEIVL 2100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10885	99.9	2146	2 Q9BZC4	Q9bzc4 homo sapien
2	10883	99.9	2146	2 Q81ZY2	Q81zy2 homo sapien
3	10881	99.9	2146	2 Q9NR73	Q9nr73 homo sapien
4	10249	94.1	2008	2 Q9GS58	Q9gs58 homo sapien
5	8431	77.4	2170	2 Q7TNJ2	Q7cnj2 rattus norv
6	8382.5	76.9	2159	2 Q91V24	Q91v24 mus musculu
7	5629	51.7	2260	2 Q8UVV4	Q8uvv4 gallus gall
8	5603.5	51.4	2261	1 ABC1_MOUSE	P41233 mus musculu
9	5593.5	51.3	2261	1 ABC1_MOUSE	Q95477 homo sapien
10	5510.5	50.6	2201	2 Q8OZB2	Q8ozb2 rattus norv
11	4904	45.0	2269	2 Q8T941	Q8t941 canis fam
12	4904	45.0	2269	2 AAR87835	Aar87835 canis fam
13	4903	45.0	2269	2 Q6T940	Q6t940 canis fam
14	4903	45.0	2269	2 AAR87836	Aar87836 canis fam
15	4902	45.0	2269	2 Q6T942	Q6t942 canis fam
16	4902	45.0	2269	2 AAR87834	Aar87834 canis fam
17	4847.5	44.5	2281	2 O02698	O02698 bos taurus
18	4819.5	44.2	2310	1 ABCR_MOUSE	O35600 mus musculu
19	4819.5	44.2	2310	2 AAH57853	Aah57853 mus muscu
20	4808	44.1	2273	1 ABCR_HUMAN	P78363 homo sapien
21	3551	32.6	2436	2 Q9HC28	Q9hc28 homo sapien
22	3541	32.5	2436	1 ABC2_HUMAN	Q9bzc7 homo sapien
23	3524.5	32.3	2434	1 ABC2_RAT	Q9ebz9 rattus norv
24	3390.5	31.1	2434	1 ABC2_MOUSE	P41234 mus musculu
25	3351.5	30.8	1771	2 Q76MW7	Q76mw7 homo sapien
26	3351.5	30.8	1771	2 BAA83014	Baa83014 homo sapi
27	3279	30.1	1487	2 Q8BPY1	Q8bpy1 mus musculu
28	2808	25.8	2595	1 AB12_HUMAN	Q86tk0 homo sapien
29	2715	24.9	1416	2 Q6ZP24	Q6zpz4 mus musculu
30	2715	24.9	1416	2 BAC98084	Bac98084 mus muscu
31	2714	24.9	2127	2 Q86W12	Q86w12 homo sapien

32 2701.5 24.8 5058 2 Q86UQ4 Q86uq4 homo sapien
33 2645.5 24.3 2143 2 Q80T20 Q80t20 mus musculu
34 2503 23.0 1704 2 BAB86781 Bab86781 homo sapi
35 2502 23.0 1704 1 ABC3_HUMAN Q99758 homo sapien
36 2471.5 22.7 1538 2 Q80XT2 Q80xt2 mus musculu
37 2468 22.7 1764 2 Q8MUA3 Q8mu3 stronglyloce
38 2463.5 22.6 1704 1 ABC3_MOUSE Q8t420 mus musculu
39 2316.5 21.3 1563 2 Q7PZY9 Q7pzy9 anopheles g
40 2272.5 20.9 1713 2 Q9VRG4 Q9vrg4 drosophila
41 2184 20.0 1679 2 Q7Q7K8 Q7q7k8 anopheles g
42 2165.5 19.9 1683 2 Q8XBG3 Q8xbg3 mus musculu
43 2165.5 19.9 1683 2 AAP73044 Aap73044 mus muscu
44 2091.5 19.2 1750 2 Q7QJG5 Q7qjg5 anopheles g
45 2070 19.0 1578 2 Q7Q5K8 Q7q5k8 anopheles g

ALIGNMENTS

RESULT 1

Q9BZC4 PRELIMINARY; PRT; 2146 AA.
AC Q9BZC4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ABC transporter member 7.
GN Name=ABCA7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21328888; PubMed=11435699;
RA Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,
RA Shulenkin S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,
RA Mattei M.G., Dean M., Denefle P., Chimini G.;
RT "Comparative analysis of the promoter structure and genomic
RT organization of the human and mouse ABCA7 gene encoding a novel ABCA
RT transporter.";
RL Cytogenet. Cell Genet. 92:264-270(2001).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF328787; AAK00959.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0006810; P:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; AAA ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234305 MW; 2391728D5AD97E75 CRC64;

Query Match 99.9%; Score 10885; DB 2; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2098; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPLEHHECHFPNKLPSAGTVPMLQGLICVNNNTCFPOLTPGEEGRSLNFDLSVSLLL 60
Db 47 PPLEHHECHFPNKLPSAGTVPMLQGLICVNNNTCFPOLTPGEEGRSLNFDLSVSLLL 106
Qy 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAPQPTKQSPLEPPMLDVALLTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAPQPTKQSPLEPPMLDVALLTSLRT 166
Qy 121 ESLGLALGQAQEPHLSLEAAEDLAQELLALSLVELRALQRPRTGTSGLLELSEALCS 180

Db 167 ESLGLALQAOEPLHSLLEAAAGDLAQELLALRLSLVELRALLQRPGTSGPLELLSEALCS 226
Qy 181 VRGSSSTVGPSLNWEASDLMLVQGPESALPDSSLSPPACSELIGALDLSHPLSRLLWRR 240
Db 227 VRGSSSTVGPSLNWEASDLMLVQGPESALPDSSLSPPACSELIGALDLSHPLSRLLWRR 286
Qy 241 LKPLILGKLLFAPDTPFTRKLMQAVNRTFEBLTLLRDVREWEMGLPRIFTFMNDSNVA 300
Db 287 LKPLILGKLLFAPDTPFTRKLMQAVNRTFEBLTLLRDVREWEMGLPRIFTFMNDSNVA 346
Qy 301 MLQRLLOWQDGRQPPGGRDHMEALRSFLDPGSGGYSMODAHADVGHVLTGRVTEC 360
Db 347 MLQRLLOWQDGRQPPGGRDHMEALRSFLDPGSGGYSMODAHADVGHVLTGRVTEC 406
Qy 361 LSLDKLEAAPSAAALVSRALQLLAHHRFWAGVFLGPEDSSDPTTEHPTDLPDGPCHVRIKI 420
Db 407 LSLDKLEAAPSAAALVSRALQLLAHHRFWAGVFLGPEDSSDPTTEHPTDLPDGPCHVRIKI 466
Qy 421 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 480
Db 467 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 526
Qy 481 LYLOQMPYPCVDDVFRLVRSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRANGL 540
Db 527 LYLOQMPYPCVDDVFRLVRSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRANGL 586
Qy 541 SRAVLWLGWFLUSCLGPFLLSAAALLVLVKLGDILPYSHPGVVFLLAFAVATVTSQFLL 600
Db 587 SRAVLWLGWFLUSCLGPFLLSAAALLVLVKLGDILPYSHPGVVFLLAFAVATVTSQFLL 646
Qy 601 SAFPSTRANLAACGLAYFSLYLPVLCVAVRDLRPAAGRVAASLLSPVAFGFCESIAL 660
Db 647 SAFPSTRANLAACGLAYFSLYLPVLCVAVRDLRPAAGRVAASLLSPVAFGFCESIAL 706
Qy 661 LEEQEGAQHNVGTRPTADVSLAQVSGLLLLDAALYGLATWLEAVCPQOYGIPEPWN 720
Db 707 LEEQEGAQHNVGTRPTADVSLAQVSGLLLLDAALYGLATWLEAVCPQOYGIPEPWN 766
Qy 721 PPFRRSWCGPRPKSPAPCPTLPDQKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 767 PPFRRSWCGPRPKSPAPCPTLPDQKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRG 826
Qy 781 LSLDFYQGHITAFIGHNGAGKTTTLLSILSGLFPSSGGSATILGHDFVSSMAAIRPHLGV 840
Db 827 LSLDFYQGHITAFIGHNGAGKTTTLLSILSGLFPSSGGSATILGHDFVSSMAAIRPHLGV 886
Qy 841 PQYNVLFDMLTVDHWFYGRKGLSAAVVGPEQDRLLQDGLVSKQSVQTRHLSGGWQR 900
Db 887 PQYNVLFDMLTVDHWFYGRKGLSAAVVGPEQDRLLQDGLVSKQSVQTRHLSGGWQR 946
Qy 901 KLSVAIAFVGSGVQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
Db 947 KLSVAIAFVGSGVQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 1006
Qy 961 RVAVVAGRLCCCSPLFLRHLSGGVYTLVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
Db 1007 RVAVVAGRLCCCSPLFLRHLSGGVYTLVKARLPLTTNEKADTMEGSDVTRQEKKN 1066
Qy 1021 GSQGSRVGTQPOLLALVQHVPGARLVBELPHELVLVLPYTGANDGSGFATLFRLEDTRLAE 1080
Db 1067 GSQGSRVGTQPOLLALVQHVPGARLVBELPHELVLVLPYTGANDGSGFATLFRLEDTRLAE 1126
Qy 1081 LRLTGYGISDTSLEBEIFLKVVVECAADTDMEDGSCGHCLCTGIAGLDVTLLRKNMPPQETA 1140
Db 1127 LRLTGYGISDTSLEBEIFLKVVVECAADTDMEDGSCGHCLCTGIAGLDVTLLRKNMPPQETA 1186
Qy 1141 LENGEPAAGSAPETDQSGPDVAGVQGWALTRQOLQALLKRLFLARRSRGLFAQLVLP 1200
Db 1187 LENGEPAAGSAPETDQSGPDVAGVQGWALTRQOLQALLKRLFLARRSRGLFAQLVLP 1246
Qy 1201 ALFVGLALVFSLLVPPFGHYPALRLSPMTYGAQVVSFFSEDAPGDPGRARLLEALLQEAGL 1260
Db 1247 ALFVGLALVFSLLVPPFGHYPALRLSPMTYGAQVVSFFSEDAPGDPGRARLLEALLQEAGL 1306

Qy 1261 EEPVQVHSSHRFSAPEVPAEVAKVILASGNWTPSPSPACQSQGARRLLPDCPAAAGGP 1320
Db 1307 EEPVQVHSSHRFSAPEVPAEVAKVILASGNWTPSPSPACQSQGARRLLPDCPAAAGGP 1366
Qy 1321 PPOAVTGSSEVVQNLTGRNLSDFLVKTYPRLVQGLKTKKWNVEVRYGGFSLGGRDPGL 1380
Db 1367 PPOAVTGSSEVVQNLTGRNLSDFLVKTYPRLVQGLKTKKWNVEVRYGGFSLGGRDPGL 1426
Qy 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSAQDSLKTIWNNKNGHWSVAF 1440
Db 1427 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSAQDSLKTIWNNKNGHWSVAF 1486
Qy 1441 VNRASNAILRAHLPDGPARRHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSVCVVFAM 1500
Db 1487 VNRASNAILRAHLPDGPARRHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSVCVVFAM 1546
Qy 1501 SFVPASFTLVLIERVTRAKHLQMGLSPTLVWLGNFMDMNCNLYLPACIVVLIFLAFO 1560
Db 1547 SFVPASFTLVLIERVTRAKHLQMGLSPTLVWLGNFMDMNCNLYLPACIVVLIFLAFO 1606
Qy 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSVSPSTAYVVLTCINLFIQINGSMAT 1620
Db 1607 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSVSPSTAYVVLTCINLFIQINGSMAT 1666
Qy 1621 FVLELPSDQKLEQVSRILKQVFLIFPHFCILGRGLIDWVRNQAMADAFERLGDQFQSPILR 1680
Db 1667 FVLELPSDQKLEQVSRILKQVFLIFPHFCILGRGLIDWVRNQAMADAFERLGDQFQSPILR 1726
Qy 1681 WEVVGNLLAMVTOGQPLFLFLLLOHRSQLLPQPRVRSPLPLGEEDEVARERERVQO 1740
Db 1727 WEVVGNLLAMVTOGQPLFLFLLLOHRSQLLPQPRVRSPLPLGEEDEVARERERVQO 1786
Qy 1741 ATQGDVLVLRNLTKVYRGQMPAVDRICLGIPECECGLLGVNGAGKTSFRMVGTGTLA 1800
Db 1787 ATQGDVLVLRNLTKVYRGQMPAVDRICLGIPECECGLLGVNGAGKTSFRMVGTGTLA 1846
Qy 1801 SRGEANVLGHSVAREPSAAHLSMGYCPQSDAIPELLTGRHLELLARLGRVPEAQVATA 1860
Db 1847 SRGEANVLGHSVAREPSAAHLSMGYCPQSDAIPELLTGRHLELLARLGRVPEAQVATA 1906
Qy 1861 GSGLARLGLSWADRPAGTYSGGNKRKLTALALVGDPAVVFDEPTTGMDDPSARRFLWN 1920
Db 1907 GSGLARLGLSWADRPAGTYSGGNKRKLTALALVGDPAVVFDEPTTGMDDPSARRFLWN 1966
Qy 1921 SLAVVREGSRVMTLTHSMECEALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLLRV 1980
Db 1967 SLAVVREGSRVMTLTHSMECEALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLLRV 2026
Qy 1981 PAARSQPAALFVAAPPGSELREAHGRLRFLPPGRCALARVFGELAVHGAHGVEDF 2040
Db 2027 PAARSQPAALFVAAPPGSELREAHGRLRFLPPGRCALARVFGELAVHGAHGVEDF 2086
Qy 2041 SVSQTMLEEVFLYFSKQDQKDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAETVL 2100
Db 2087 SVSQTMLEEVFLYFSKQDQKDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAETVL 2146

RESULT 2

Q81ZV2 PRELIMINARY; PRT; 2146 AA.
AC Q81ZV2, 2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ABC transporter ABCA7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20549028; PubMed=11095984;

Db 1127 LRLTGYGSDTSLBEIFLKVVVECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1186
 Qy 1141 LENGEPAAGSAPETDQSGPDAGVGVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
 Db 1187 LENGEPAAGSAPETDQSGPDAGVGVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1246
 Qy 1201 ALFVGLALVFLSLIVPPFGHYPALRLSPTMYGAQVSFFSEADPGDGRARLLEALLQAGL 1260
 Db 1247 ALFVGLALVFLSLIVPPFGHYPALRLSPTMYGAQVSFFSEADPGDGRARLLEALLQAGL 1306
 Qy 1261 EEPVQVHSSHRFSAPEVAEYAKVLAGSNMTPESPACQSQPGARRLLPDPCPAAGGP 1320
 Db 1307 EEPVQVHSSHRFSAPEVAEYAKVLAGSNMTPESPACQSQPGARRLLPDPCPAAGGP 1366
 Qy 1321 PPOQAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEYRGFSLGGRDPGL 1380
 Db 1367 PPOQAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEYRGFSLGGRDPGL 1426
 Qy 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLLDAQDSLKTFWNNKGWHSWAF 1440
 Db 1427 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLLDAQDSLKTFWNNKGWHSWAF 1486
 Qy 1441 VNRSNAILRAHLPPGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSIQVVFAM 1500
 Db 1487 VNRSNAILRAHLPPGPARHAHSITTLNHPNLTKQELFEAALMASSVDVLVSIQVVFAM 1546
 Qy 1501 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWGNFLMDMCMNYLPACIVVLIFLAQ 1560
 Db 1547 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWGNFLMDMCMNYLPACIVVLIFLAQ 1606
 Qy 1561 QRAVVAIPANIPALLLLLLLGVMSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMAT 1620
 Db 1607 QRAVVAIPANIPALLLLLLLGVMSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMAT 1666
 Qy 1621 FVLELFDQKLQEVSRILKQVFLFPFHFLGRGLIDMVRNQAMADAFERLGDROFQSPLR 1680
 Db 1667 FVLELFDQKLQEVSRILKQVFLFPFHFLGRGLIDMVRNQAMADAFERLGDROFQSPLR 1726
 Qy 1681 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQPRVRSPLPGSEEDVARERERVQ 1740
 Db 1727 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQPRVRSPLPGSEEDVARERERVQ 1786
 Qy 1741 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPGCEGGLGVNGAGAKTSTFRMTVGTDLA 1800
 Db 1787 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPGCEGGLGVNGAGAKTSTFRMTVGTDLA 1846
 Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCFQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1860
 Db 1847 SRGEAVLAGHSVAREPSAAHLSMGYCFQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1906
 Qy 1861 GSGLARLGLSWYADRPAGTYSGGNKRKLATALLVGDPAVVFLDEPTTGMDPSARRELWN 1920
 Db 1907 GSGLARLGLSWYADRPAGTYSGGNKRKLATALLVGDPAVVFLDEPTTGMDPSARRELWN 1966
 Qy 1921 SLLAVVREGSVMLTSSMECEALCSRLATMNGRFRCLGSPHLLKGRFAAGHTLTLRV 1980
 Db 1967 SLLAVVREGSVMLTSSMECEALCSRLATMNGRFRCLGSPHLLKGRFAAGHTLTLRV 2026
 Qy 1981 PAARSQPAAPVAAEFPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAESHGVEDF 2040
 Db 2027 PAARSQPAAPVAAEFPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAESHGVEDF 2086
 Qy 2041 SVSOTMLEEVLVYPSKQDKDEDETEOKEAGVGVDPAPGLOHPRKVSQFLLDDPSTAETVL 2100
 Db 2087 SVSOTMLEEVLVYPSKQDKDEDETEOKEAGVGVDPAPGLOHPRKVSQFLLDDPSTAETVL 2146

RESULT 3
 Q9NR73 PRELIMINARY; PRT; 2146 AA.
 AC Q9NR73
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Macrophage ABC transporter.
 GN Name=ABCA7;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20334305; PubMed=10873640;
 RA Kaminski W.E., Orso E., Diederich W., Klucken J., Drobnik W.,
 RA Schmitz G.;
 RT "Identification of a novel human sterol-sensitive ATP-binding cassette
 transporter (ABCA7).";
 RL Biochem. Biophys. Res. Commun. 273:532-538(2000).
 CC -/- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AF250238; AF85794.1; -.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0005524; F:ATP binding; TAS.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; TAS.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding.
 SQ SEQUENCE 2146 AA; 234468 MW; 679B16EB2D75FF0D CRC64;
 Query Match 99.9%; Score 10881; DB 2; Length 2146;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2098; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PPLEHECHFPNKLPSAGTVPWLQGLICNVNNTCPQLTPGPEPGRSLNFDLSRL 60
 Db 47 PPLEHECHFPNKLPSAGTVPWLQGLICNVNNTCPQLTPGPEPGRSLNFDLSRL 106
 Qy 61 ADARTVLGGASARTLAGLKLATLRAARSTAAQPTKOSPLEPMLDVAELLTSLRT 120
 Db 107 ADARTVLGGASARTLAGLKLATLRAARSTAAQPTKOSPLEPMLDVAELLTSLRT 166
 Qy 121 ESLGLALGOAEPHLSLEAAEDLAQELLALRSIVELRALLQPRGTSGLLELSEALCS 180
 Db 167 ESLGLALGOAEPHLSLEAAEDLAQELLALRSIVELRALLQPRGTSGLLELSEALCS 226
 Qy 181 VRGFSSTVGPSSLNWEASDLMELVQEPESALPDSSLPACSELIGALDGHPLSLWRR 240
 Db 227 VRGFSSTVGPSSLNWEASDLMELVQEPESALPDSSLPACSELIGALDGHPLSLWRR 286
 Qy 241 LKPLILGLKLLFADPTPTFKLMAQVARTPEELTLRDVREVWMLGPRIETFMNDSNVA 300
 Db 287 LKPLILGLKLLFADPTPTFKLMAQVARTPEELTLRDVREVWMLGPRIETFMNDSNVA 346
 Qy 301 MLORLQMQDEGRQRPGGRDHEALRSFLDPGSGYSGWQDAHADVGHVLVGLTVTEC 360
 Db 347 MLORLQMQDEGRQRPGGRDHEALRSFLDPGSGYSGWQDAHADVGHVLVGLTVTEC 406
 Qy 361 LSLDKLEAAPSEAAALYSRALQLLAHRFWAGVVLGPESSDPTEHPTDLPDGHVRIKI 420
 Db 407 LSLDKLEAAPSEAAALYSRALQLLAHRFWAGVVLGPESSDPTEHPTDLPDGHVRIKI 466
 Qy 421 RMDIDVVRTNKLDRDFWDGPAADPLTDLYVWGGFVYLODLVERAAVVLGSANPRAG 480
 Db 467 RMDIDVVRTNKLDRDFWDGPAADPLTDLYVWGGFVYLODLVERAAVVLGSANPRAG 526
 Qy 481 LYLQOMPYPYVDDVFLRSLPLFLTLAWYSVTLTKAVVREKETRLDTRMAMGL 540
 Db 527 LYLQOMPYPYVDDVFLRSLPLFLTLAWYSVTLTKAVVREKETRLDTRMAMGL 586
 Qy 541 SRAVLMLGWFLSCIGPFLISAALLVLKLDILPYSHPGVWFLFAAFVAVTQSFLL 600
 Db 587 SRAVLMLGWFLSCIGPFLISAALLVLKLDILPYSHPGVWFLFAAFVAVTQSFLL 646

QY 601 SAFFSRANLAAACGGGLAYFSLYLPVYLCVAMDRLPAGGRVAASLLSPVAFPGCESLAL 660
DB 647 SAFFSRANLAAACGGGLAYFSLYLPVYLCVAMDRLPAGGRVAASLLSPVAFPGCESLAL 706
QY 661 LEEQEGAGHNVGTRPTADVFSLAQVSGLLLDAAALGLATWYLEAVCPQGYGIPEPWN 720
DB 707 LEEQEGAGHNVGTRPTADVFSLAQVSGLLLDAAALGLATWYLEAVCPQGYGIPEPWN 766
QY 721 FFRRSYWCPCPPKSPACPTPLDPKVLVEBAPGLSPGVSRSLKEKFPSPALRG 780
DB 767 FFRRSYWCPCPPKSPACPTPLDPKVLVEBAPGLSPGVSRSLKEKFPSPALRG 826
QY 781 LSLDFYQGHITAPLGHNGAGKTTTILSLGLPPSPGGSFALIGHDVSRSSMAAIRPHLGYC 840
DB 827 LSLDFYQGHITAPLGHNGAGKTTTILSLGLPPSPGGSFALIGHDVSRSSMAAIRPHLGYC 886
QY 841 PQYNVLFDMLTVDHVFYGRILKGLSAAVGPEDRLQDVLGVSKQSVQTRHLSGMQR 900
DB 887 PQYNVLFDMLTVDHVFYGRILKGLSAAVGPEDRLQDVLGVSKQSVQTRHLSGMQR 946
QY 901 KLSVAIAFGVGSQVWILDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAELLGD 960
DB 947 KLSVAIAFGVGSQVWILDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAELLGD 1006
QY 961 RVAVVAGRLCCCGSPLFLRRHLGSGYLLTVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
DB 1007 RVAVVAGRLCCCGSPLFLRRHLGSGYLLTVKARLPLTTNEKADTMEGSDVTRQEKKN 1066
QY 1021 GSQSRVGTQQLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATLFRLELTRLAE 1080
DB 1067 GSQSRVGTQQLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATLFRLELTRLAE 1126
QY 1081 LRLTGVSIDTSEELFLKVEECADTDMEGSCGHLCTGIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGVSIDTSEELFLKVEECADTDMEGSCGHLCTGIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAGSAPETDQSGSDVAVGRVQGWALTQQQLALLKFLARRSRRLGFAQIVLP 1200
DB 1187 LENGEPAGSAPETDQSGSDVAVGRVQGWALTQQQLALLKFLARRSRRLGFAQIVLP 1246
QY 1201 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSPFSDADPGDGRARLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSPFSDADPGDGRARLLEALLQEAGL 1306
QY 1261 BEPPVQSHRSFAPEVPAEVAKVILASGNWTPESPACQCSQPCARLLPDCPAAAGP 1320
DB 1307 BEPPVQSHRSFAPEVPAEVAKVILASGNWTPESPACQCSQPCARLLPDCPAAAGP 1366
QY 1321 PPQAVTGSVEVVQNLTRNLSDFLVKTYPRLVROGLTKKWVNEVRYGGFSLGGRDPGL 1380
DB 1367 PPQAVTGSVEVVQNLTRNLSDFLVKTYPRLVROGLTKKWVNEVRYGGFSLGGRDPGL 1426
QY 1381 PSGQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIFWNNKGWHSWAF 1440
DB 1427 PSGQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIFWNNKGWHSWAF 1486
QY 1441 VNRSASNAILRAHLPGRPARHAHSITTLNHPNLNTEQLESEALMASSVDVLVSIQVWAM 1500
DB 1487 VNRSASNAILRAHLPGRPARHAHSITTLNHPNLNTEQLESEALMASSVDVLVSIQVWAM 1546
QY 1501 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPACIVLLIFLAFO 1560
DB 1547 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPACIVLLIFLAFO 1606
QY 1561 QRAYVAPANLPALELLLYGWSITPLMYPASFPFSPSTAYVTLTINLFTINGSMAT 1620
DB 1607 QRAYVAPANLPALELLLYGWSITPLMYPASFPFSPSTAYVTLTINLFTINGSMAT 1666
QY 1621 FVLELFSQKLOEVSRIILKQVFLIIPPHCLGRGLIDMVRNQAMADAFERLGRDQSQPLR 1680
DB 1667 FVLELFSQKLOEVSRIILKQVFLIIPPHCLGRGLIDMVRNQAMADAFERLGRDQSQPLR 1726

QY 1681 NEVVGKNLLAMVIOGPLELLTLLQHRSQLLPQPRVSLPLLGEDEEDVARERERVQ 1740
DB 1727 NEVVGKNLLAMVIOGPLELLTLLQHRSQLLPQPRVSLPLLGEDEEDVARERERVQ 1786
QY 1741 ATQGDVLVLRNLTKYRGORMPAVDRLCLGIPGECFGLLVGNVAGKTTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLVLRNLTKYRGORMPAVDRLCLGIPGECFGLLVGNVAGKTTSTFRMVTGDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOAQT 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOAQT 1906
QY 1861 GSGLARLGLSWYADRPAGTYSGNKRKLATALALVGDPAVFLDPTTGMDFSARRFLWN 1920
DB 1907 GSGLARLGLSWYADRPAGTYSGNKRKLATALALVGDPAVFLDPTTGMDFSARRFLWN 1966
QY 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
DB 1967 SLLAVVREGRSVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQAAAAFVAABFPGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAHGVDF 2040
DB 2027 PAARSQAAAAFVAABFPGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAHGVDF 2086
QY 2041 SVSQTMLBEVFLYFSKQDQKDETEEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2100
DB 2087 SVSQTMLBEVFLYFSKQDQKDETEEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 4
Q96S58 PRELIMINARY; PRT; 2008 AA.
AC Q96S58;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
DE ABCA-SSN.
GN Name=ABCA7/ABCA-SSN;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21255283; PubMed=11355874;
RA Tanaka A., Ikeda Y., Abe-Dohmae S., Arakawa R., Sadanami K.,
RA Kidera A., Nakagawa S., Nagase T., Aoki R., Kioka N., Amachi T.,
RA Yokoyama S., Ueda K.;
RT "Human ABCA1 Contains a Large Amino-Terminal Extracellular Domain
Homologous to an Epitope of Sjogren's Syndrome";
RL Biochem. Biophys. Res. Commun. 283:1019-1025(2001).
CC -|- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB055390; BAB62294.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0001656; F:nucleotide binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003313; ABC binding.
DR Pfam; PF00005; ABC tran.;
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2008 AA; 218616 MW; 226FF85C24230B90 CRC64;

Query Match 94.1%; Score 10249; DB 2; Length 2008;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	121	ESIGLALGOAEPPLHSLLEAAEDLAEALSLVELRALLORPRTSGPGLLELSEALCS	180
Db	29	ESLGLALGOAEPPLHSLLEAAEDLAEALSLVELRALLORPRTSGPGLLELSEALCS	88
Qy	181	VRGSSSTVGPLSNWYEASDLMELVQGPESALPDSSLSPPACSELIGALDHSPLSRLLWRR	240
Db	89	VRGSSSTVGPLSNWYEASDLMELVQGPESALPDSSLSPPACSELIGALDHSPLSRLLWRR	148
Qy	241	LKPLILGKLLPAPDTPTRKLAQVNRTEFELTLRLDVRVEMWMLGPRIFTFMNDSSNVA	300
Db	149	LKPLILGKLLPAPDTPTRKLAQVNRTEFELTLRLDVRVEMWMLGPRIFTFMNDSSNVA	208
Qy	301	MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVCHLVGTIGRVTETC	360
Db	209	MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVCHLVGTIGRVTETC	268
Qy	361	LSLDKLEAAPSEALVSRALQLLAHRFWAGVFLGPEDESDPTHEHTPDLPQGHVRIKI	420
Db	269	LSLDKLEAAPSEALVSRALQLLAHRFWAGVFLGPEDESDPTHEHTPDLPQGHVRIKI	328
Qy	421	RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	480
Db	329	RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	388
Qy	481	LYLOQMPYPCVDDVFLRVLRSRSLPLFLTLAWIYSVTLTKVAVREKETRLDTRAMGL	540
Db	389	LYLOQMPYPCVDDVFLRVLRSRSLPLFLTLAWIYSVTLTKVAVREKETRLDTRAMGL	448
Qy	541	SRAVLWLGWFLSLCLGPFLLSAAALVLVLKLGDIILPYSHPGVVFLLAFAFVATVTSQFLLL	600
Db	449	SRAVLWLGWFLSLCLGPFLLSAAALVLVLKLGDIILPYSHPGVVFLLAFAFVATVTSQFLLL	508
Qy	601	SAFPSRANLAAACGGLAYFSLYLYLVLCVWRDLRDPAGGRVAASLLSPVAFPGCESLAL	660
Db	509	SAFPSRANLAAACGGLAYFSLYLYLVLCVWRDLRDPAGGRVAASLLSPVAFPGCESLAL	568
Qy	661	LEEGEGAHNVNCTRTADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQOYGIPEPWN	720
Db	569	LEEGEGAHNVNCTRTADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQOYGIPEPWN	628
Qy	721	PPFRSRYWCGPRPKSPAPCTPLDPKVLVEEAPPGSLSPGVSVRSLEKRPFGSPQPALRG	780
Db	629	PPFRSRYWCGPRPKSPAPCTPLDPKVLVEEAPPGSLSPGVSVRSLEKRPFGSPQPALRG	688
Qy	781	LSLDFYQGHITAFILGHNGAGKTTLSLSGLFPSPGSGSAFTILGHDRSSMAAIRPHLGVC	840
Db	689	LSLDFYQGHITAFILGHNGAGKTTLSLSGLFPSPGSGSAFTILGHDRSSMAAIRPHLGVC	748
Qy	841	PQYNVLFDMLTVDHVMFYGLKGLSAAVVGPEODRLLQDVGLYSKOSVQTRHLSGGMQR	900
Db	749	PQYNVLFDMLTVDHVMFYGLKGLSAAVVGPEODRLLQDVGLYSKOSVQTRHLSGGMQR	808
Qy	901	KLVAIAFVGGQVQVILDEPTAGVDPASRRGIIWELLKLYREGRTLILSTHLDABELLGD	960
Db	809	KLVAIAFVGGQVQVILDEPTAGVDPASRRGIIWELLKLYREGRTLILSTHLDABELLGD	868
Qy	961	RVAVVAGGRCLCCGSPFLFRHRHLSGGYLLTLVKARLPLTTNEKADTDMEGSVDTROBKN	1020
Db	869	RVAVVAGGRCLCCGSPFLFRHRHLSGGYLLTLVKARLPLTTNEKADTDMEGSVDTROBKN	928
Qy	1021	GSQSRVGTTPQLLALVQHVPGARLVEELPHELVLVLPYTCAGHDSFATLRFRELDTRLAE	1080
Db	929	GSQSRVGTTPQLLALVQHVPGARLVEELPHELVLVLPYTCAGHDSFATLRFRELDTRLAE	988
Qy	1081	LRLTYGISTDSLEEIFLKVVVECAADTDMEDSCGQHLCGTGIAGLDVTLLRKMPPOETA	1140
Db	989	LRLTYGISTDSLEEIFLKVVVECAADTDMEDSCGQHLCGTGIAGLDVTLLRKMPPOETA	1048
Qy	1141	LENCEPAGSAPETDQGGPDVAVRQGWALTRQOLQALLKRLFLARRSRRLGPAQIVLP	1200
Db	1049	LENCEPAGSAPETDQGGPDVAVRQGWALTRQOLQALLKRLFLARRSRRLGPAQIVLP	1108

Qy	1201	ALFVGLALVFSLVPPFGHPALRLSPTMYGAOVSPFSEDAPGDPGRARLLEALLQEAGL	1260
Db	1109	ALFVGLALVFSLVPPFGHPALRLSPTMYGAOVSPFSEDAPGDPGRARLLEALLQEAGL	1168
Qy	1261	EEPPVQHSRRFSAPEVAEPAVKYASGNWTPSPSPACQSQPGARRLLPDCPAAAGGP	1320
Db	1169	EEPPVQHSRRFSAPEVAEPAVKYASGNWTPSPSPACQSQPGARRLLPDCPAAAGGP	1228
Qy	1321	PPQAVTGSSEVVQNLTRNLSDFLVKTYPRLYRQGLKTKKWNVEYRGFSLGGRPGL	1380
Db	1229	PPQAVTGSSEVVQNLTRNLSDFLVKTYPRLYRQGLKTKKWNVEYRGFSLGGRPGL	1288
Qy	1381	PSQELGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSILDAQSLKTLFWNNKGWHSVAF	1440
Db	1289	PSQELGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSILDAQSLKTLFWNNKGWHSVAF	1348
Qy	1441	VNRASNAILRAHLPFGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSIQVVFAM	1500
Db	1349	VNRASNAILRAHLPFGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSIQVVFAM	1408
Qy	1501	SFVPASFTLVLIIEBRTVTRAKHLQIMGGLSPTLWLGNFMDMCNLYLPACIVVLIFLAFQ	1560
Db	1409	SFVPASFTLVLIIEBRTVTRAKHLQIMGGLSPTLWLGNFMDMCNLYLPACIVVLIFLAFQ	1468
Qy	1561	QRAYVAPANLPALLLLLLYXWSITPLMYPASPFSSVPSTAYVVLTCINLFIGINGSMAT	1620
Db	1469	QRAYVAPANLPALLLLLLYXWSITPLMYPASPFSSVPSTAYVVLTCINLFIGINGSMAT	1528
Qy	1621	FVLELFSQDKLOEVSRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLR	1680
Db	1529	FVLELFSQDKLOEVSRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLR	1588
Qy	1681	WEVVGKLLAMWIOGPELFTLLQHRSQLLPQPRVRSPLIGEEDEDAVARERVVQG	1740
Db	1589	WEVVGKLLAMWIOGPELFTLLQHRSQLLPQPRVRSPLIGEEDEDAVARERVVQG	1648
Qy	1741	ATQGDVLVRLNLTQVYRGORMPAVDRLCIGIPGCECFGLLVNGAGKTSIFRMVGTDLA	1800
Db	1649	ATQGDVLVRLNLTQVYRGORMPAVDRLCIGIPGCECFGLLVNGAGKTSIFRMVGTDLA	1708
Qy	1801	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVAQTA	1860
Db	1709	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVAQTA	1768
Qy	1861	GSGLARLGLSWYADRPAGTYSNGNKRKLATALALVGDPAVVFLEDEPTTGMDDPSARRLWN	1920
Db	1769	GSGLARLGLSWYADRPAGTYSNGNKRKLATALALVGDPAVVFLEDEPTTGMDDPSARRLWN	1828
Qy	1921	SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRPRCLGSPHLLKGRFAAGHTLTLRV	1980
Db	1829	SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRPRCLGSPHLLKGRFAAGHTLTLRV	1888
Qy	1981	PAARSQPAALFAAEFFPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHGVEDF	2040
Db	1889	PAARSQPAALFAAEFFPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHGVEDF	1948
Qy	2041	SVSGTMLLEEVFLYPSKQDQKDETEOEKAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL	2100
Db	1949	SVSGTMLLEEVFLYPSKQDQKDETEOEKAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL	2008

RESULT 5
Q7TNJ2 ID Q7TNJ2 PRELIMINARY; PRT; 2170 AA.
AC Q7TNJ2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATP-binding cassette transporter sub-family A member 7.
GN Name=ABCA7;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Rattus.

```

NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
STRAIN=Wister; TISSUE=platelet;
Sasaki M., Nada S., Yamaguchi A.;
Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the ABC transporter family.
CC EMBL; AB097814; BAC81426.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004009; F:ATP binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:000610; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding.
SQ SEQUENCE 2170 AA; 237718 MW; 003C8DF70B8744CE CRC64;

Query Match 77.4%; Score 8431; DB 2; Length 2170;
Best Local Similarity 76.5%; Pred. No. 0;
Matches 1646; Conserved 154; Mismatches 272; Indels 80; Gaps 12;

QY 1 PLEHHECHFPNKPPLPSAGTVPMWQGLICNVNNTCPQPTQCGEEGRSLNENDSLVRL 60
DB 1 PLEHHECHFPNKPPLPSAGTVPMWQGLICNVNNTCPQPTQCGEEGRSLNENDSLVRL 60
QY 61 ADARTVLGGASAHRTIAGLGKLIATLRAARSTAQP-----QPTKQSPLEPMLDVAELLTS 116
DB 107 ADAHTVLGGHSTQDMLAAGKLPVLRVAGSGANPQESNQPAKQG-----SVTELEK 159
QY 117 LRTESLGLALCOAPEPLHSLLEAEDLAQELALRLSVLVELRALLQRPRTSGPLELSE 176
DB 160 ILQRASLETVLGQAQDSMEKFSDATRTVAQELLTLPSLVELRALLRRPRGSAGSELISE 219
QY 177 ALCSVRGSPSTVGPSLNWYEASDLMLVGOEPESALPOSSLSPACSELIGALDSPLSRL 236
DB 220 ALCSKTPSSPGSLNWTYEAQINEFMGPFLAPTLPOSSLSPACSEFVGLDDHPVSLR 279
QY 237 LWRLKPLILGKLLFAPDTFTFKLMAQVNRFTFEELTLRLDREYVWMLGPRIFTFMNDS 296
DB 280 LWRLKPLILGKLLFAPDTFTFKLMAQVNRFTFEELTLRLDREYVWMLGPRIFTFMNDS 339
QY 297 SNVAMLQRLQMDGRRQPRCGRDHMEALRSFLDPGSGGYSDQADHADVGHVLTGLR 356
DB 340 TNVAMLQKLLDVEGTGWQQTQKQKQLEAIRDPLPSRGRYNNQEAHADMGRLAEILQG 399
QY 357 VTECLSLDKLEAAPSEALVSRALQALLAHRRFWAGVVFGLGPDSDPTHEHTP-DLGRGH 415
DB 400 ILCEVSLDKLEAVPSEALVSRALGERRRRLWAGIVFLSPHPHLDSEPPSTTTGPGH 459
QY 416 VRIKIRMDIDVTRINKIRDREDFWDPGAADPLTLRYVWGGFVYLQDLVRAAVRVLGA 475
DB 460 LRVKIRMDIDVTRINKIRDREDFWDPGADPLMDLRYVWGGFVYLQDLLEQAAVRLSGR 519
QY 476 NPARGLYLOQMPYPCVDDVFLRVLSRLPLFLTLAWIYSVTLTKAVVREKETRLRDTM 535
DB 520 DSRAGLYLQMPHPYPCVDDVFLRVLSRLPLFLTLAWIYSVTLTKAVVREKETRLRDTM 579
QY 536 RAMGLSRVAVLWGLFSLCLGPELLSAALLVVLKGLDILPYSHPGVWFLFLAAFAVATV 595
DB 580 RAMGLSRVAVLWGLFSLCLGPELLSAALLVVLKGLDILPYSHPGVWFLFLAAFAVATV 639
QY 596 QSFLLSAFFSRANLAAACGLAYFSLYLPYVLCVAVRDLRAGGRVAAASLLSPVAFGFC 655
DB 640 QSFLLSAFFSRANLAAACGLAYFALYLPYVLCVAVRDLRAGGRVAAASLLSPVAFGFC 699
QY 656 ESLALLEEGGGAQWHNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLSAVCPGQYGI 715
DB 1760 LAMVAQGFLLITLLQLHNRNLLPQPKSRLLPPLGDEEDDVREVRERVTGATGCDVLV 1819
```

QY 1749 LRNI.TKYRGORP.VA.DRL.CL.GIPPGECFGLGVNGAGKTSTFRMTGDTLASRGEAVLA 1808
 Db 1820 LRDI.TKYRGORS.PAVDHLCLGIPPGECFGLGVNGAGKTSTFRMTGDTLPSSGEAVLA 1879
 QY 1809 GHVAREPSAAHLSMGYCPQSDAIFELLTGREHLEILARLGRVPEAQVATAGSLARLG 1868
 Db 1880 GHVQAQPSAAHRSMGYCPQSDAIFDLITGREHLELFARLGRVPEAQVATALSGLVRLG 1939
 QY 1869 LSWYADRPAGTYSGNKRK.LAT.ALA.LVGD.PAVVFLDEPTTGMPSAARFLWNSLLAVRE 1928
 Db 1940 LPSYADRPAGTYSGNKRK.LAT.ALA.LVGD.PAVVFLDEPTTGMPSAARFLWNSLLSVRE 1999
 QY 1929 GRSVLT.SHSMECEAL.CSR.LA.I.MVNGR.FRCLGSPQHLKGRFAAGHTLT.LRVPAARSQPA 1988
 Db 2000 GRSVLT.SHSMECEAL.CSR.LA.I.MVNGR.FRCLGSAQHLKSRFGAGHTLT.LRVPPDQPEPA 2059
 QY 1989 AAFVAAEPFSGSELREAHGGR.LR.FOL.PPGRCALARV.FGELAVHGAEGHGVDFSVSQTMLE 2048
 Db 2060 IAFVTTPDAELREVHGS.RLR.FOL.PPGCGCT.LARV.FRE.LAAQKAGHGVDFSVSQTILE 2119
 QY 2049 EVFLYFSKDGKDBDTEBEQ.EAGVGVDPAP.FGLQHPKRVSVQFLDDPSTAETVL 2100
 Db 2120 EVFLYFSKDGQEBEGSQETETREVS.TFGLQHPKRVSVFLEDPSVSETVI 2170
 RESULT 6
 Q91V24
 ID Q91V24 PRELIMINARY; PRT; 2159 AA.
 AC Q91V24;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE ATP-binding cassette transporter sub-family A member 7.
 GN Name=Abca7;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=O1129, and DBA/2;
 RX MEDLINE=21328888; PubMed=11435699;
 RA Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,
 RA Shulenin S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,
 RA Mattei M.G., Dean M., Denefle P., Chimini G.;
 RT "Comparative analysis of the promoter structure and genomic
 RT organization of the human and mouse ABCA7 gene encoding a novel ABCA
 RT transporter".
 RL Cytogenet. Cell Genet. 92:264-270(2001).
 CC -!- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AF287142; AAK56863.1; -;
 DR EMBL; AF287141; AAK56862.1; -;
 DR HSSP; P58301; 1F2U.
 DR MGD; MGI:1351646; Abca7.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:000524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA_2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 KW ATP-binding.
 SQ SEQUENCE 2159 AA; 236881 MW; CD2BE3FE0DB822B CRC64;

Query Match 76.9%; Score 8382.5; DB 2; Length 2159;

Best Local Similarity 76.4%; Pred. No. 0;
 Matches 1635; Conservative 160; Mismatches 277; Indels 69; Gaps 11;
 QY 1 PPLEHHECHFPNKP.LPSAGTVP.WLQGLICNVNNTCFPQ.LTPGEEPRGLSNFNDLSLRLL 60
 Db 47 PPLEHHECHFPNKP.LPSAGTVP.WLQGLVCNVNNSCFQHFTPGKPGVLSNFKDSL.SRLL 106
 QY 61 ADARTVLGASAHRTIAGLKGKLIAT.AAARSTAQP.---OPTKQSPLEPPMLDVAELLTS 116
 Db 107 ADRTVLVGHHSIQDMLDALGK.LIPVURAVGGGARPOESDQPTSQG.-----SVTKLLEK 159
 QY 117 LRTESIGLALGAQOAP.LHSLLEAEADLAQELIALRS.LVELRALLQRPRTSGPLELSE 176
 Db 160 ILQASLD.PVLGAQASMRKFS.DAIRDLAQELLT.LPSLMELRALLRRPGSAGSLELVSE 219
 QY 177 ALCVSGPSTVTPGSLNWTVEASDLMELVGOEPESAL.PDSLSLPACSELGALDSHPLSRL 236
 Db 220 ALCSTKGPSPGSLNWTVEANQLNBFMGPEVAPAL.PDNSLSPACSEFVGTLDHDPVSRLL 279
 QY 237 LWRRLKPLILGKLLFAPDTPFTK.LMAQVNRTEELTLRLDRVREVMELGPRIFTFMNDS 296
 Db 280 LWRRLKPLILGKLLFAPDTPFTK.LMAQVNRTEELTLRLDRVREVMELGPRIFTFMNDS 339
 QY 297 SNVAM.LQRLLQ.MODEGRQRP.PGCRDHMEALRS.FLDPGSGYQSWQDAHADVGHVLTGR 356
 Db 340 TNVAM.LQRLLD.VGGTGQRQOT.PRAQKLEA.IKDFLDP.SRGGYSWREAHADMGRLAGILGQ 399
 QY 357 VTECLSLDKLEAAP.SAALVSRA.LQ.LLAH.RFWAGVVF.LGPESSDPTTEHPTDLPGRHV 416
 Db 400 MMECVSLDKLEAVP.SBEALVSRA.LELIGERR.LWAGIVFL.SPEHPLDPSELSPALSPGHL 459
 QY 417 RIKIRMDIDVVTNKT.RDRFWD.PGPAADPLT.DLRYVMGFFVYLQDLVERAAVRVLSGAN 476
 Db 460 RPKIRMDIDVVTNKT.RDKFDPG.PSAPFMDLRYVMGFFVYLQDLLEQAAVRVLSGGN 519
 QY 477 PRAGLYLQOMPYP.CYVDDVFL.RVLSRSLPLFLT.LAMIYSVTLTKAVREKETRLRDTMR 536
 Db 520 SRTGLYLQMPHE.CYVDDVFL.RVLSRSLPLFLT.LAMIYSVTLTKAVREKETRLRDTMR 579
 QY 537 AMGLSRAVLWLGW.FLSC.LGFLLSAALLVLVLKGDILPYSHPGVVF.LFAAFATVATQ 596
 Db 580 AMGLSRAVLWLGW.FLSC.LGFLLSAALLVLVLKGNILPYSHPWVFLFLFAAFATVATQ 639
 QY 597 SFTLSAFFSRANLAAACGGI.AVSLV.PYVLCVAVDRDL.PAGGRVAASLLSPVAFGGE 656
 Db 640 SFTLSAFFSRANLAAACGGI.AV.FALVY.PVLCVAVRERHLGGLLAASLLSPVAFGGE 699
 QY 657 SLALLBEEQSGAQW.HNVGTRPTAD.VFSLAQVSGLLLLD.ALYGLATWYLEAVCPGYGIP 716
 Db 700 SLALLBEEQSGAQW.HNLGTGPAEDV.FSLAQVSAFLLLD.VIYGLALWYLEAVCPGYGIP 759
 QY 717 EPWNFPFRRSYWC.GPRPPKSPAP.CPTPLDPKVLVEAPPGLS.PGVSVRSLEKPPGSPQP 776
 Db 760 EPWNFPFRRSYWC.GPGPKPSVLAPAPQDPKVLVEEPPPLGLVP.GVSRIGLKGHFRGCPQP 819
 QY 777 ALRGLSLDFYQGHITAF.LGHNGAGKTTTIL.SGLPSPSGSAFILGHDRVSSMAARPH 836
 Db 820 ALQGLNLDFYEGHITAF.LGHNGAGKTTTIL.SGLPSPSGSASILGHQVQTNNAARPH 879
 QY 837 LGVCPQVNVLFDM.LTVDH.VWFYGR.LKGLSAAVVGPEQDRLLOQVGLVSVKQSVQTRHLSG 896
 Db 880 LGICPQVNVLFDM.LTVEEH.VWFYGR.LKGVSAAMGPERERLIRDVGLTLKARDIQTRHLSG 939
 QY 897 GMQRKLSVAIFAVGGSQVILDEPTAGVD.PASRRGIWELLKLYREGRTLLSTHHLDEAE 956
 Db 940 GMQRKLSVAIFAVGGSRVIMDEPTAGVD.PASRRGIWELLKLYREGRTLLSTHHLDEAE 999
 QY 957 LLGDRVAVVAGGR.LCCGSP.LFRR.HL.GSGVYLT.LVKARLPLTNE.-KATDMEGSDVTR 1015
 Db 1000 LLGDRVAVVAGGR.LCCGSP.LFRR.HL.GCGYLT.LVKSSQSLVTHDKAGDSE.-----DPR 1054
 QY 1016 QEKKNGSQSGS-----RVGTPQLLALVQHWPVPGARLVEELP 1050

```
Db 1055 REKSDGNGRTSDTATRTGTSKSNQAPAGVPIPTPSTARILELVQHVPGQAQVEDLP 1114
Qy 1051 HELVLVLPYTGADGDSFATLRELDRLAELRLTYGIGSDTSLBETFLKVVBECAADTDM 1110
Db 1115 HELLVLPYAGALDGSFAMVFOELQOELLEGLTYGIGSDTNLEIFLKVVED--AHREG 1172
Qy 1111 EDCSGQHICTGIAGLDVTLRLKMPPEQETALENGEPAGSAPETDQSGPDVAVRGVGNAL 1170
Db 1173 GDSRQQLHRT-----CTPQPTGPESVLENGELA-----PQGLAPNA-AQVQGWTL 1219
Qy 1171 TFOQLQALLKFLARSRGLFAQIVLPALFVGLALVFSLIVPPFGHPYALRLSPMY 1230
Db 1220 TCQQRALLHKFLARSRGLFAQIVLPALFVGLALVFSLIVPPFGHPYALRLSPMY 1279
Qy 1231 GAQVSFFSDAPCDPGARLLLEALQEALEPPVQVHSHR-----FSAPEVPA 1279
Db 1280 GPQVSFFSADAPGDPNRMKLEALLGEAGLQPSQDKDARSECTHSLACYFTTPEVPP 1339
Qy 1280 EVAKVLASGNWTPSPSPACQSPGARLLPDCPAAAGGPPPPQAVTSGSGBVGNLTGR 1339
Db 1340 DVASILASGNWTPSPSPACQSPGARLLPDCPAGAGGPPPPQAVAGLGSVGNLTGR 1399
Qy 1340 NLSDFLVKTYPLVRQGLKTKWNEVRVYGGFSLGGRDPLGSLGQELGSRVSEELWALLSP 1399
Db 1400 NVSDFLVKTYPLVRQGLKTKWNEVRVYGGFSLGGRDPLGSLGQELGSRVSEELWALLSP 1459
Qy 1400 LPGAALDRVLKMLTAWHSLDAQDSLKIWNFNKGHSMVAVFNASNAITLRAHLPPGPAP 1459
Db 1460 QFGNALDRILNLTQWALGLDARNLSKIWNFNKGHSMVAVFNANRANGLLHALLSPGVR 1519
Qy 1460 HAHSITTLNHPNLTKQEQLSEALMASSVDVLVSVICVVPFAMSFVPASFTLVLIBERVTRA 1519
Db 1520 HAHSITTLNHPNLTKQEQLSEALMASSVDVLVSVICVVPFAMSFVPASFTLVLIBERVTRA 1579
Qy 1520 KHLQMGSLPTLYMGLNPLDMCNVLPACIVLIFLAFQORAYVAPANLPALLLILL 1579
Db 1580 KHLQVSLGPOTLYMGLNPLDMCNVLPACIVLIFLAFQORAYVAPANLPALLLILL 1639
Qy 1580 YGWSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMATVLELFSQDLQVSRILK 1639
Db 1640 YGWSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMATVLELFSQDLQVSRILK 1699
Qy 1640 QVFLIPHPCLGRGLIDMVNRQAMADAFERLGDROFQSPRLRWEVGNKLLAMVIOGLPL 1699
Db 1700 QVFLIPHPCLGRGLIDMVNRQAMADAFERLGDROFQSPRLRWDIIIGKLLAMVIOGLPL 1759
Qy 1700 LFTLLQHRSLLPQVRSLPLICEDEEDVARERERVVOGATQGDVLVRLTKVYRGQ 1759
Db 1760 LFTLLQHRSLLPQVRSLPLICEDEEDVARERERVVOGATQGDVLVRLTKVYRGQ 1819
Qy 1760 RMPVDRLCIGIPPGCEGGLGVNGAGKTSTFRMTGDTLASRGAVALAGHSVAREPSAA 1819
Db 1820 RMPVDRLCIGIPPGCEGGLGVNGAGKTSTFRMTGDTLASRGAVALAGHSVAREPSAA 1879
Qy 1820 HLSMGYCQSDAIPELLTGREHLELLARLGVPEAQVATAGSGLARLGLSWYADRPAQT 1879
Db 1880 HRSMGYCQSDAIPELLTGREHLELLARLGVPEAQVATAGSGLARLGLSWYADRPAQT 1939
Qy 1880 YSGGNKRLATALVGDPAVFLDEPTGMDPSARRFLWNSLLAVVREGSRVMTLSHM 1939
Db 1940 YSGGNKRLATALVGDPAVFLDEPTGMDPSARRFLWNSLLAVVREGSRVMTLSHM 1999
Qy 1940 ECEALCSRLAIWNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQAPAAVFAAEFFGS 1999
Db 2000 ECEALCSRLAIWNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQAPAAVFAAEFFGS 2059
Qy 2000 ELREAAGGRRLRQLPPGRCALARYFVELAHLVGAHGVDEDFSVSOTMLEEFLYFSKQOG 2059
Db 2060 ELREAAGGRRLRQLPPGRCALARYFVELAHLVGAHGVDEDFSVSOTMLEEFLYFSKQOG 2119
Qy 2060 KDEDTEEQEAGVGDVDPAGLQHPKRVSQFLDDPDDTAETVL 2100
Db 2120 EBEESRQ--EABEEBVSXPGRQHPKRVSRFLBDDPSSVETMI 2159
```

```
RESULT 7
Q8UVV4
ID Q8UVV4 PRELIMINARY; PRT; 2260 AA.
AC Q8UVV4;
DT 01-MAR-2002 (TREMREL. 20, Created)
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)
DE 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE ATP-binding cassette transporter 1.
GN Name=ASCAI;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RP Attie A.D., Brooks-Wilson A.R., Walker D., McManus B.,
RA Gray-Kellar M.P., MacDonald M.L.E., Roomp K., Tebon A., Zhang L.-H.,
RA Mulligan J., Sensesen C., Bitgood J.J., Cook M.E., Kastelein J.J.P.,
RA Hayden M.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF362377; AAL56247.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti.; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2260 AA; 254070 MW; 19D137F342F98662 CRC64;
```

```
Query Match 51.7%; Score 5629; DB 2; Length 2260;
Best Local Similarity 49.8%; Pred. No. 0;
Matches 1117; Conservative 363; Mismatches 588; Indels 174; Gaps 25;
Qy 1 PPLEHHECHFPKPLPSAGTVPLQGLICNVNNTCFPOLTPGEERGLSNFNDLSVRL 60
Db 47 PPYQHQHCHFPNKAMPAGTLPWIGLIICNANNPCFRTPTGESPVGIVGNFASVRL 106
Qy 61 ADARTVLGASAHRTLAGLGLIATLRAA----- 89
Db 107 SDAKRLLYSQDTSIKDVQKVLAKRLKGNSSGLDLKRLDPLVDNETFSDPLRHVSNP 166
Qy 90 -----RSTAQPOPTKQSPLEPP 106
Db 167 SSAAVEELDAEVLNQLKQIVSGYRIQLRDLNCSSALSEFTLIQNRSVAMDSEAFCLTPKE 226
Qy 107 MLDVAEL-----LTSLLTESLGLALGOAQLHLLAAEDLAQELALRLSLVELR 158
Db 227 TLHAAELAFRANLPLKPLQREIFFNSLRDLSETVEALRDSLGLKVLKELLSKWSNDR 286
Qy 159 ---ALLQPRGTSGPLEL---LSEALCSVRGFSSTVGPSLNWYEASDLMELVG----QSP 208
Db 287 QEVMFLLTNVNASSTQIQAVSRIVCGHPGGGLKIKSLNWDNNYKALFGGNSTEDD 346
Qy 209 ESALPDSSLSPACSELIGALDHPRLSLRLWRRLKPLILGLKLLFADPTFTKRLMAQVNR 268
Db 347 VTFNFDNSTTTCYNELMKNLESSPLRIIRWALPLKPLIGLVKLYTPDTPAIRKINAEVNR 406
Qy 269 FEELTLARDVREVMELGPRIFTFMDSSNVAMLQRLIQ-----MODEGR 313
Db 407 FOELGVFRDLGGMWEEISPKIWTFWESSQEMDLIRTLKSKALMDLHLHPASNTWTVAR 466
Qy 314 ---RQPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLEAAP 370
```

Db 467 FLKHP-----EFEADNGWV-----YTWDAFNETDRAITQISRMECVNLDKLEPVA 515
Qy 371 SEALVSRALQLLAHRPWAGVFLGPDSDPTEHPPTDLPDGPCHVIRKIRMDIDVVTRT 430
Db 516 TEVLRLINKSLLEDDERRFWAGVF--TEIADNTELPQ-----HVYKIRMDIDNVERT 567
Qy 431 NKIBDRWDPCPADPLTDLRYVNGGFVYLQDLVERAAVRVLSGANPRAGILQOOMPYC 490
Db 568 NKIKDGVNDPGRADPEDMKYVWGFTYLDQVVEQAIIRVQTEKKTGYVVOOMPYC 627
Qy 491 YVDDVFLVLSRSLPLFTLTAWISVTLTCAVYREKETRLDRTRAMGLSRVILWLGWF 550
Db 628 YVDDIFLVRMSRNPFLTWLAWISVAVIIGIYVEKEARLKETWRIMGLDNGILNLSWF 687
Qy 551 LSCIGPFLLSAALVLVLKGDIIIPYSHPGVVFLLAFAVATVQSFLLSAPFSRANLA 610
Db 688 ISSLIPLMSAGLLVLILKMGNLLPYSDPSVVFVFLSIFGIVTILQCLISTVFSRANLA 747
Qy 611 AACGLAVFSLYLPVLCVARDRLPAGGRVAASLLSPVAFGFCESIALLEEAGEAOW 670
Db 748 AACGGIVFTLYLPLVLCVARDRLPAGGRVAASLLSPVAFGFCESIALLEEAGEAOW 807
Qy 671 HNVGTRP--TADVSLAQVSGILLDDAALYGLATWYLEAVCPGOYGIPEPNFPRRSYWC 729
Db 808 DNPFPESLEEDGFSITTSANVMDLFTLYGWVWYIESVFPQYGI PRPYFPFKSYWP 867
Qy 730 GPRPCKSPACPTPLDP--KVLVEBAPGLSPGVSVRSLEKRFPGSPQALRGSLDFYQG 788
Db 868 GEESQDRQLHPDQKQSPVCEKPEMHLISLGVSIQNLVKVYRDGKVAVDGLTLNFEY 927
Qy 789 HITAFHNGAGKTTLSILSGLEPPSGGSFALIGHDVRSSMAIRHLGVCPOYNVLF 848
Db 928 QITSFLHNGAGKTTLSILSGLEPPSGGSFALIGHDVRSSMAIRHLGVCPOYNVLF 987
Qy 849 MLTVDHWFYGRUKLGLSAAVGPEDRLQDVLGVSKQSVQTRHLGGQKRLSVAIAF 908
Db 988 LITVEEHIWFARLKGIPKVKKEEMEQAMDVLPHKLKARTSKLSGGQKRLSVALAF 1047
Qy 909 VGSQVILDEPTAGVDPASRRGIWELLKYREGRTILILSTHLDDEALLGDRVAVVAG 968
Db 1048 VGSQVILDEPTAGVDPASRRGIWELLKYREGRTILILSTHLDDEALLGDRVAVVAG 1107
Qy 969 RLCCGSPFLRRLHLSGYLITLTKARLPLTN-----EKADTDMESVDTRQ 1016
Db 1108 KLCCVGSFLKNGQLGTGYLTLVKDQVDSLSRNSSTSVYLKDDSVSSSSDAGL 1167
Qy 1017 EKNGSQSRVGTQPLLALVOHWVPGARLVEELPHELVLVLPYTGADHSGFATLFLRLDT 1076
Db 1168 GSDHESDITLIDVSAISNLITKHVPPEARLVEDIGHELTYVLVYKAAKEGAFVELFHEID 1227
Qy 1077 RLAEIRLTGYSISTSLBEEIFLKVBECAADTDMEDSCGHLCTGIAGLDVTLRLKMP 1136
Db 1228 RLSDLGISYISGETTLEEIFLKVADSDGVDASDGTLPARRNRRAFG--DRQSLRPFT 1286
Qy 1137 QETALE--NCEPAGSAPETDQSGPDVAG--RVQWALTQOOLQALLKPELLARRSRG 1192
Db 1287 EDDAFDNDSDIDPESRETDLSDMGQSGYQMKGKLSQOQFWALLKRLITAKSRKG 1346
Qy 1193 LFAQIVLPALFVGLVFLVPPFGHYPALRISPTMYGAQVSPFSDADGDPGRARLLE 1252
Db 1347 FFAQIVLPALFVGLVFLVPPFGHYPALRISPTMYGAQVSPFSDADGDPGRARLLE 1406
Qy 1253 ALLQEQAG-----LEBPVQHSRHSFSAPEVPAEVAKVASGNWTPESPSPACQS 1302
Db 1407 ALLNKPFGTRCMQGHISIDPTCTVGQKEWTTASVDPDVSLEIL--RGNWSMENPSPCECS 1465
Qy 1303 QPGARRLLPCPAAGGPPPOAVTGSVEVONLTGNSLDFLYKTYTFLVRQGLTKKW 1362
Db 1466 NEKTKMLPVCPPGAGGLPPQREODTADILQNLITGNSIDLYLTKYTAQIIGSLKKNKI 1525
Qy 1363 VNEVRYGFGSLGGDRDPG--LPSGQBLGRSVEELMALLSPLPGCALDRVLKNTAWAHSIDA 1421

Db 1526 VNEFRYGGFSLGARSHVLPSPSNEVTDIAIKQVKKILELAQSGSGDRFLNNLASFMKGLDT 1585
Qy 1422 QDSLKIWFNNKGHSMVAFVNRASNAILLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEA 1481
Db 1586 KXNVKVFNNKGHSMVAFVNRASNAILLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEA 1645
Qy 1482 ALMASSVDVLVSICVVFAMSFVPASFTVLVIERVTRAKHLQMLGGLSPTLYWLNFLWD 1541
Db 1646 ALMTTSDVLVSICVVFAMSFVPASFTVLVIERVTRAKHLQMLGGLSPTLYWLNFLWD 1705
Qy 1542 MCNLYVPACTVILFLAFQOQRAYVAPANLALLLLLYLWGSITPLMPYSPFFSVPSTA 1601
Db 1706 MCNLYVPACTVILFLAFQOQRAYVAPANLALLLLLYLWGSITPLMPYSPFFSVPSTA 1765
Qy 1602 YVVLTCINLBIGINGSMAFVLELFDQDKLQEVSRILKQVFLFPHPHFCGLRGLIDWVRN 1661
Db 1766 YVVLTSYNLFIGINGSVATFVLELFTNNKLNININDILKSVFLFPHPHFCGLRGLIDWVRN 1825
Qy 1662 AMADAFERLGDROFQSPLRWEVVGKLLAWIOGPIFLFLFTLLLOHRSOLLPOVRSLP 1721
Db 1826 AMADAFERLGDROFQSPLRWEVVGKLLAWIOGPIFLFLFTLLLOHRSOLLPOVRSLP 1885
Qy 1722 LLGEEDVVARERERVQCATQGDVLVLRNLTKVYRQRMVAVDRCLGIPPGCECFLLG 1781
Db 1886 PVNDEDEDVNRERQRIISGGQSDILEIRELTKIYRMKRKPAVDRCVGPCECFLLG 1945
Qy 1782 VNGAGKTSIPRWVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREH 1841
Db 1946 VNGAGKTSIPRWVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREH 2005
Qy 1842 LEILARLGVPEAQVTAQTAGSGLARGLSMYADRPAGTVGGNKRKLATALLVGDPAV 1901
Db 2006 LEFALLRGLVPEAQVTAQTAGSGLARGLSMYADRPAGTVGGNKRKLATALLVGDPAV 2065
Qy 1902 FLDEPTTGMDPSARRLFWNSLLAVVREGSRVMLTSHSMECECALCSRLAMVNGRFRCLG 1961
Db 2066 FLDEPTTGMDPSARRLFWNSLLAVVREGSRVMLTSHSMECECALCSRLAMVNGRFRCLG 2125
Qy 1962 SPQHLKGRFAAGHTLTLRVPAAARS--QPAAFVAAEPFPGSELREAHGGRURFOLPGGRC 2019
Db 2126 SVQHLKGRFAGDGTIVVRIAGGNPDLPKPVVEFFGHAFPGSVLKEKERNMLQYL--PSSQS 2184
Qy 2020 ALARVGEALVHCAEHGVEDFSVQTMLEVFYFSDQGDQEDTDE--QKEAGVGVD 2076
Db 2185 SLARISVLSQNKRLHIEDYSVQTTLDQVFNFAQDQSDDDHTDKLSLHKNQTV--VDI 2243
Qy 2077 APGLQHPKRVSOFLDDPSTAET 2098
Db 2244 A-----ILNSFLQDEKVKES 2258
PRT; 2261 AA.
AC P41233;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette transporter 1) (ATP-binding cassette 1) (ABC-1).
GN Name=Abca1; Synonyms=Abcl1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Macrophage;
RX MEDLINE=94375008; PubMed=8086782;
RA Luciani M.F., Denizot F., Savary S., Mattei M.-G., Chimini G.;
RT "Cloning of two novel ABC transporters mapping on human chromosome 9."
RT 9."
RL Genomics 21:150-159(1994).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J;
RC MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RA Qiu Y., Cavelier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RX "Human and mouse ABCA1 comparative sequencing and transgenesis studies
RT revealing novel regulatory sequences.";
RL Genomics 73:66-76(2001).
RN [3]
RP INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not
RL ABCG1 via an LXR-independent pathway.";
RJ J. Lipid Res. 43:952-959(2002).
RN [4]
RP DOWN-REGULATION BY ENDOTOXIN
RX PubMed=12777468; DOI=10.1194/jlr.M300100-JLR200;
RA Khovidhunkit W., Moser A.H., Shigenaga J.K., Grunfeld C.,
RT "Endotoxin down-regulates ABCG5 and ABCG8 in mouse liver and ABCA1 and
RJ ABCG1 in J774 murine macrophages: differential role of LXR.";
RL J. Lipid Res. 44:1728-1736(2003).
CC -!- FUNCTION: CAMP-dependent and sulfonyleurea-sensitive anion
CC transporter. Key gatekeeper influencing intracellular cholesterol
CC transport (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed in adult tissues. Highest
CC levels are found in pregnant uterus and uterus.
CC -!- INDUCTION: Down-regulated by endotoxins (LPS) or cytokines (TNF
CC and IL-1) in J774 macrophages. The down-regulation by endotoxin in
CC macrophages is not likely to be mediated by the liver X
CC receptor/retinoic X receptor (LXR/RXR).
CC -!- DOMAIN: Multifunctional polypeptide with two homologous halves,
CC each containing an hydrophobic membrane-anchoring domain and an
CC ATP binding cassette (ABC) domain.
CC -!- PTM: Phosphorylation on Ser-2054 regulates phospholipid efflux (By
CC similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; X75926; CA453530.1; ALT_INIT.
CC EMBL; AF287263; AAG39073.1; ALT_INIT.
CC MGD; MGI:99607; Abca1.
CC GO; GO:0008203; P:cholesterol metabolism; IDA.
CC GO; GO:0030301; P:cholesterol transport; IDA.
CC GO; GO:0042158; P:lipoprotein biosynthesis; IMP.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 2.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Phosphorylation; Transmembrane; Transport.
FT TRANSMEM 26 42 Potential.
FT TRANSMEM 640 656 Potential.
FT TRANSMEM 690 706 Potential.
FT TRANSMEM 717 733 Potential.
FT TRANSMEM 749 765 Potential.
FT TRANSMEM 771 787 Potential.
FT TRANSMEM 1041 1057 Potential.
FT TRANSMEM 1351 1367 Potential.
FT TRANSMEM 1661 1677 Potential.
FT TRANSMEM 1708 1724 Potential.
FT TRANSMEM 1737 1753 Potential.
FT TRANSMEM 1775 1791 Potential.
FT TRANSMEM 1854 1870 Potential.
FT NP_BIND 933 940 ATP (Potential).

FT NP_BIND 1946 1953 ATP (Potential).
FT MOD_RES 1042 1042 Phosphoserine (by PKA) (By similarity).
FT MOD_RES 2054 2054 Phosphoserine (by PKA) (By similarity).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 98 98 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 161 161 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 244 244 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 292 292 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 337 337 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 349 349 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 478 478 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 489 489 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 820 820 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1144 1144 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1294 1294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1453 1453 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1499 1499 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1504 1504 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1637 1637 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2044 2044 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2238 2238 N-linked (GlcNAc...) (Potential).
FT CONFLICT 1567 1568 Missing (in Ref. 2).
FT CONFLICT 2024 2024 Missing (in Ref. 2).
SQ SEQUENCE 2261 AA, 253972 MW, F0C3C5F1CFE34F1 CRC64;

Query Match 51.4%; Score 5603.5; DB 1; Length 2261;
Best Local Similarity 49.6%; Pred. No. 0;
Matches 1118; Conservative 363; Mismatches 577; Indels 195; Gaps 27;

QY 1 PPLEHHECHFNKPLPSAGTVPWLOGLICVNNNTCFPQITCEEGRLSNFNDSLSRL 60
DB 47 PPYEQHECHFNKAMPASGTLFWVGQIICNANNPCFRYPTEGAPGVGNFKSVSR 106
QY 61 ADARTVLGGASAHRTLAGLGLKIATRAAR-----STAQOQ 96
DB 107 SDAQRLLYSQDTSIKDMHKVLMRLQIKHPNSNKLQDFLVNDNTPSGFLQHLNLS 166
QY 97 PTKOSPLBPPM-----LDVAE 112
DB 167 STVDSLLQANVGLQVFLQVQLGLHSLCNGSKLEEIIQLGDAEVSALCGLPRKLL 226
QY 113 LL-----TSLRTESLGLAQOEPHLSLEAEADLAQELLALRSLV 155
DB 227 RVLRYNMDILKPVVTKLNSGTLPTQHIA-----EATVLLDLSGLAQELFTKWS 279
QY 156 ELR---ALLQPRGTSGPLEL---LSEALCSVRGSPSTVGPSLNWYEASDLMELVG--- 205
DB 280 DMRQEVMLTNVNSSSTQIYQAVSRVCGHPEGGLKIKSLNWNEDNNYKALFGGNT 339
QY 206 QEPESALPDSSLPACSELIGALDSHPLSRLRLRLKPLILGLKLFAPDTPPTRKLM 265
DB 340 EEDVDTFYDNGSTTPYCNLMKNLESSPLSRIIWKALKELLVGLKILYTDTPATRV 399
QY 266 NRTTEELTLDRVREWMELCPRIETFNDSNVNMLQRILOMQ--DGRQRPRGGRDH 323
DB 400 NKTQELAVFHDLEGMEELSPOIWTFMENSQEMDLVRLTLLDSRGNDQFWEQKLG 459
QY 324 MEALRSFL-----DPGSGGYSWODAHADVHLVGLGRVTECLSDKLEAAPEAL 375
DB 460 AQDINAFIAKNPVEDVQSPNGSVYTWREAFNETNQAIQIISFMECVNLKLEPTE 519
QY 376 VSRALQLLAEHRFWAGVVFGLGPDSSDPTEHTPDLG--PGHVRIRKIMDIDVVR 433
DB 520 INKSMELLDERKFWAGIVFTG-----ITPDSVELPHKVKYKIRMDIDNVRTNK 569
QY 434 RDRFWDPGPADPITDLRYVWGGFYVLODLVERAAVRVLSGANPRAGILYQMPFCY 493
DB 570 KDGWYDPGRADPFDMRYWGGFYAYLQDVVEQAIIRVLTSSEKTKGVYVQMPFCY 629

Qy	494	DVFLRVLSRSLPLPLTLTAWIYSVTTLTKVAVVREKETRLDTRMANGLSRAVLWLGWFLSC	553
Db	630	DIFLVRNRSWPLPMTLAWIYSVAVITKSIIVYEKEARLKETRMITGINDGNIGLWFSWFVS	689
Qy	554	LGPFLLSAALLVVLKGLDILPYSHPGVWFLFLAAPAVATVTQSFLLSAFSPSRANLAAAC	613
Db	690	LIPULVAGLLWLTKLGNLLPYSDPSVFWFVSFAMVTLQCLFLLSTLPSRANLAAAC	749
Qy	614	GGLAYFLYLPPYLVCVNRDLRPAAGGRVAASLLSPVAFGFCESLALLEBQSGCAQWHNV	673
Db	750	GGIYFTLYLPPYLVCVQDYGFSIKIPASLLSPVAFGFCVYALFESQIGVGQWDL	809
Qy	674	GTRPT-ADVFSLAQVSGLLLDAAALYGLATWYIEAVCPQGVGPIPEPWNFPFRSRSYWCGPR	732
Db	810	FESPEVEDGFNLTTAVSNMLPDTFLYGWMTWYIEAVFPQGVGIPRPWYFPCTKSYWFGEE	869
Qy	733	PPKSPACPPTLD-PKVLVEEAPGLSPGVSVRSLEKFPGSPQPALRGLSLDFYQGHIT	791
Db	870	IDEKSHPGSQKGSEICMEEPFHLGLVSIQNLVKYRDGMKVAVDGALANFYEGQIT	929
Qy	792	AFLCHNGAGKTTTILSGLFPPSPGGSAFILGHVDRSSMAIRPHLCVCPQYNVLFDMLT	851
Db	930	SFLGHNGAGKTTTMSIILTLGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCQHNVLFDMLT	989
Qy	852	VDEHVWYFRLKGLSAAVWGPEDRLQDYL-VSKOSVQTRHLSGQMOKLSVAIAFVG	910
Db	990	VEEHIWFYARLKGLSEKHVAEMQALDVLGPPSKLKSKTQLSGQMOKLSVALAFVG	1049
Qy	911	GSQVVLDEPTAGVDPASRGIWELLKLYBREGRTLILSTHLDABEALLGDRVAVVAGGRL	970
Db	1050	GSKVVLDEPTAGVDPSRRGIWELLKLYQGRITILSTHMDADTLGRIAIISHGKL	1109
Qy	971	CCGSSPLFRRHLGSGYYLTLVKARLPLTTN-----EKADTMEGVSVDTRQEK	1018
Db	1110	CCVGSSFLKNQJGTGYLTLVLKXDVESSLSSCNRSSSTVCLCKEDSVSQSSDAGLGS	1169
Qy	1019	KNGSQSRVGTQQALLALVQHWWPGARLVEELPHELVLVLYPTGAHDGSAFRLFELDTRL	1078
Db	1170	DHESDTLTIDVSAISNLIRKHVSERLVEDIGHELTVYLPYEAKEGAFVELPHEIDRL	1229
Qy	1079	ABELRTGYGSDTSLBEIFLKVVECAADTDMEG-----SCGHLCTGI	1123
Db	1230	SDLGISSYGISTLEBIFLKVABESGVDAETSDGTLPARNRRAFGDKQSC-LHPPTED	1288
Qy	1124	AGLDVTLRLKMPPOETALENGEPAGSAPEITDQSGSDPAVG--RVQGWALTRQQLQALLK	1181
Db	1289	DAVD-----PNDSID-----PESRETDLLSGMDKGSGYQJUKWKLTQQQFVALLWK	1335
Qy	1182	RFLIARRSRRLFAQIVLPALFVGLALVFSLLVPPFGHYPALRLSPTWYGAQVSFFSEDA	1241
Db	1336	RLLIARRSRKGFQIAQIVLPAVFCIALVFSLLVPPFGKYPSELEQPMWYNEQITFVSNDA	1395
Qy	1242	PGDPRARLLEALLQEAG-----LSEPPVQHSHRFSAPEVPAEVAKVLASGNWT	1291
Db	1396	PEDMGTQELLNALTQDPGFGRTRCMEGNPIDPTPCLAGEEOWTISPVQPSIVDLFQNGNWT	1455
Qy	1292	PESPSPACQSPQCARLLPDCPAAAGPPPPQAVTGSGEVQNLTCGNLSDFELVTKYPR	1351
Db	1456	MKNPSPACQSSDKIKKMLPVCPPGAGGLPPQKQKTAIDLQMLTCGNLSDYLVKTYVQ	1515
Qy	1352	LVROGLTKKWNVRYRGFSLG-GRDPGLPSGOELGRSVEELWALLSPLPGGALDRVLK	1410
Db	1516	IIAKSLKNKIWNBFRRGGFSLGVSNSQALPPPSHEVNDAIKQMKLLKLTQDISADRFLS	1575
Qy	1411	NLTAWAHSLLDAQDSLKTWFNNKGWHSWVAFVNRASNAILRAHLPPGPARHAHSTITLNH	1470
Db	1576	SLGRFMAGLDPTKNVVKVFNFNKGWHAISSEFLNVINNALRANLQKGNPQSYGITAENHP	1635
Qy	1471	LNLTKEOLSEALMASSVDVLVSTCVVFPMSFVPASTFLVLIBERVYTRAKHLQMLGSLP	1530
Db	1636	LNLTKQGLSEVALMTTSDVDLVSTCVLFPMSFVPASFPVFLVLIBERVYTRAKHLQIPISG	1695
Qy	1531	TYWLGFLWDMCNLYLPACIVLILFLAFOORAYVAPANIPALLLLLLLYGWSITPLMYP	1590

Db	1696	VIYLSNFVWDMCNVVPATLVIIIFCPOKSYSVSTNLPVIALLLLVGWSITPLMYP	1755
Qy	1591	ASPFSSVPSTAYVVLTCINLFIGINGSMAFVLELFSDDQKLOEVSRIKQVFLIFPHFCL	1650
Db	1756	ASFVKIPSTAYVVLVSIVNLFIGINGSVATFVLELFTNNKLANDINDILKSVFLIFPHFCL	1815
Qy	1651	GRGLIDMVRNQAMADAFERLGDQFQSPURWEVVGKXLLAMWJTGQFLFLFTLLQLHRSQ	1710
Db	1816	GRGLIDMVKNQAMADALEREFGENRFVSPISWDLVVGRNLFAMAVEGVVFLITVLIOVRFF	1875
Qy	1711	LLPQPRVRSIPLIGEEDVDVAREERVVQATOGDVLVLRNLTKVYRGORMPAVDRICLG	1770
Db	1876	IRPRPVKAKLPPLNDEDEDVRRERQRILDDGGQNDILEIKELTKIYRRKKRPVADRICIG	1935
Qy	1771	IPGECFGLLVGNAGAGKTSFRMVTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPQSD	1830
Db	1936	IPGECFGLLVGNAGAGKSTTFKMLTGDPVTRGDGAFLNKNSILNTHVHQNGYCFQPD	1995
Qy	1831	AIPELLTGRHEHLLARLGRVPEAQVATAGSGLARLGLSWYADRPAQTVSGNGKXKLAT	1890
Db	1996	AIPELLTGRHEHVEFFALLRGVPEKEVGKFGEMAIKGLGVKEKYASNSGNGKXKLST	2055
Qy	1891	ALALVGDPAVVFLEDEPTTGMDDPSARFELNLSLLAVRGESVMLTSHSMECEALCSRLA	1950
Db	2056	AMALIGGPPVVFLEDEPTTGMDDPKARFELNWCALLSVKEGSSVVLTSHSMECEALCTRMA	2115
Qy	1951	IMVNGRPRCLGSPQHLKLGRAAGHTLTLRVPAARS--QPAFAAFVAAEFPGSELREAHGGR	2008
Db	2116	IMVNGRPRCLGSPVQHLNKRFGDGTIVVRIAGSNPDLKPVOEFFGLAFPGSVLKEKXNM	2175
Qy	2009	LRFQLPGGRCALARVFGELAVHGAHGVEDFSVSQTMLEEVLYFSKQGGKDE---DTE	2065
Db	2176	LQYQL-PSSLSLARIFSILSQSKRLHIEDYSVSQTTLLDQVFNFAKQSDDDHLKDL	2234
Qy	2066	EQKEAGVGVDPAQGLQHPKXVSOLFDDPSTAET	2098
Db	2235	LHKNQTV-VDVAV-----LTSFLODEKVKES	2259
RESULT 9			
ABCL_HUMAN			
ID	ABCL_HUMAN	STANDARD;	PRT; 2261 AA.
AC	Q95477; Q96856; Q9GT85; Q9NQV4; Q9UN06; Q9UN07; Q9UN08; Q9UN09;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux regulatory protein).		
GN	Name=ABCA1; Synonyms=ABCA1, CERP;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20345099; PubMed=10884428;		
RA	Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,		
RA	Freeman L.A., Cheng J.-F., Osorio J., Remaley A.T., Yang X.-P.,		
RA	Haudenschild C.C., Prades C., Chimini G., Blackman E.E.,		
RA	Francois T.L., Duverger N., Rubin E.M., Rosier M., Deneffe P.,		
RA	Frederickson D.S., Brewer H.B. Jr.;		
RT	"Complete genomic sequence of the human ABCA1 gene: analysis of the		
RL	human and mouse ATP-binding cassette A promoter.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Skin;		
RA	Schwartz K., Lawn R.M., Wade D.P.;		
RT	"ABCA1 gene expression and apoA-I-mediated cholesterol efflux are		
RL	regulated by LXR.";		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		


```
Db 1773 VNLFTGINGSVATFLEIFTDNKNANNINDILKSVLEIFPHFCLGRGLDMVKNQAMADAL 1832
Qy 1668 ERLGDRQFOSPLRWVGVGNLLAMVIOGFLPLFTLLQHRSQLLPQPRVSLPLGREED 1727
Db 1833 ERFGENRFVPSLWDLVGNLFAMAVEGVVFLIIVLQYRFFIRPRVNAKLSPLNDED 1892
Qy 1728 EDVABERERVQAGATQGVVLVRLNLTQYVQRMQMPAVDRCLCIGIPPGCEFCGLLVNGAGK 1787
Db 1893 EDVRRERQIILDDGGONDILEIKELTKIYRRKRKPAVDRIICVIGIPPGCEFCGLLVNGAGK 1952
Qy 1798 TSTFMVMTGDTLASGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGBHEHLELLAR 1847
Db 1953 STTFKMLTGDFTVTRGDAPLANKNSLNIHEVQNMGYCPOQDAITELLTGREHVEFFAL 2012
Qy 1848 LRGVPEAQVAOTAGSGLARLGLSVADRPAGTYSGNKRKLATALLVGDPAVFLDEPT 1907
Db 2013 LRGVPEKEVGKGEWAIRKGLGLVKYGEKYGAGNYSGNKRKLSTAWALIGGPVFLDEPT 2072
Qy 1908 TGMDSARRFLWNSLLAVVREGSRVWLTSHSMBECCALCSRLAIMVNGRFRCLGSPQHLK 1967
Db 2073 TGMDPKARRFLWNCALSVVKEGRSVVLTSHSMBECCALCTRMALMVNGRFRCLGSPQHLK 2132
Qy 1968 GRPAGHTILTILRVPAARS--QPAAPFAVAAEPGSELREAHGRLRFQLPFGRCMARVF 2025
Db 2133 NRGDGYTIVVRIAGSNPLDKPVODFFGLAPFGSVPEKXHRNMLQYQL-PSLSLSLARIF 2191
Qy 2026 GELAVHGAEBHGVEDFSVOTMLEEVLFPYSDQKDE--DTEEOKEAGVGDPAQGLQH 2082
Db 2192 SILSOSKRLHIEDVSVQTLQDVFNPAKQSDDDHLKULSLHKNQTV-VDVAV----- 2246
Qy 2083 PKRVSQFLDDPSTABT 2098
Db 2247 ---LTSFLQDEKVKES 2259

RESULT 10
Q80ZB2
ID Q80ZB2 PRELIMINARY; PRT; 2201 AA.
AC Q80ZB2;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE ATP-binding cassette 1.
GN Name=Abcal;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Ananthanarayanan M.; Mirza M.F.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY208182; AA033557.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; P:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS0211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2201 AA; 246553 MW; B1472978BFC3E6B8 CRC64;
```

Query Match 50.6%; Score 5510.5; DB 2; Length 2201;
Best Local Similarity 49.5%; Pred. No. 0;
Matches 1108; Conservative 361; Mismatches 575; Indels 195; Gaps 28;

```
Qy 15 LPSAGTVPWLOGLICNVNNTCFPOLTPCEEPGRLSNFNDLSVRSLLADARTVLGCASAH 74
Db 1 MPSAGTLFWVGIIICNANPCFRPTPGEAFOGVGNFNKSVSRLSFSDAQRIILLYSQKDT 60
Qy 75 TLAGLGLKIATLRAAR-----STAQPQPTKQSPLE----- 104
Db 61 SIRDHMKVLRTLQOIKHNSNKLQDPLVDNTEFSGFLQHSLSLPRSAVDNLLQADVSLQ 120
Qy 105 -----PPMLDVAEILL----- 114
Db 121 KVFLQGVOLHIASLNGSKLEBIIIRPDLKVSALCSLPREKLDAPERELRSNMDILKPMV 180
Qy 115 -----TSLLRTESLGLALGOAQEPLHSLLEAAEDLAQELLALRSLSLVEALLQ-----RP 164
Db 181 TKLANSTSLPTQHLA-----EATTLLDSLGLAQELFSTKSWDNROQVEMFLTNNV 233
Qy 165 RGTSGPL-ELLSEALCSVRGSPSTVGPSLNWYEASDLMLVGV---QBPESALPDSSISP 219
Db 234 SGSSTQIYQAVSRIVCGHPEGGLKIKSLNWNENNYKALFGGNGTEEDTDTFYDNSTTP 293
Qy 220 ACSELI GALDSDHPLSRLLWRBLKPLILGLKLPADPTPTKLMQAVNPTFEELTLRDVR 279
Db 294 YCNLDMKNLESPLSRIIWKALKPLLIKILYPTDTPATROQVMAEVNKTFOELALFPDL 353
Qy 280 EVWEMLGPRIETFWMNDSNVAMLQRLQMO--DEGRORPPGGRDHMEALRSFL----- 331
Db 354 GMBELSPQIWTWFESSQEMDLVRPMLDLRGNDQFWEKLDGLYTAQDIWAPLAKNPED 413
Qy 332 --DFSGGYSQMDAHADVGLVGTGRVTECLSLDLEAAPSEALVSRALQLLAHRFW 389
Db 414 VQSPNGSVYTWEAFNETNQAIQTSRFMECVNLNLEPIPTVTLINKSMDDLDAKFS 473
Qy 390 AGVVFGLGPEDSSDTEHPTDGLG--PGHVRIRKIMDIDVVTTRNKIRDRPNDPGAADPL 447
Db 474 AGIDFTG-----ITPDSVELPHHVOKIRMDIDNVERTNKIKDGYWDPGRADPF 523
Qy 448 TDLRVNMGGVYVQLQDLVERAAVRVLSGANPRAGVLQOMPYPVYVDDVFLVLSRSLPLF 507
Db 524 EDMRYVNGGFAYLQDVVEQAIIRVLTGTEKTYVQOMPYPVYVDDVFLVMSRSMPLF 583
Qy 508 LTLAWISVTLTVKAVVREKETRLRDTMRAMGLSRAVLWGLWFLSCGLGPFLLSAAALLV 567
Db 584 MTLAWISVAVIIKSIIVVEKARLKTMRIMGLDNGILWFSWFISSLLPLLVSAAGLVII 643
Qy 568 LKLGDIILPYSHPGVYVFLAFAVATVTSPLLSAFPSRANLAAACGLAFVSLYLPVYL 627
Db 644 LKLGDIILPYSDPSVYVFLSVFAVVTILQCEFLISTLFSRXNLAACGGIIFTLYLPVYL 703
Qy 628 CVAWRDLRIPAGGRVAAISLLSPVAFGFCESLALLEEOGEOGAQWNVGRPT-ADVFSLAQ 686
Db 704 CVAQDVVGSIKIFASLLSPVAFGFCGYPALFEEOGIGVQWNLFPKSPVEEDGFNLTT 763
Qy 687 VSGLLLLDAALYGLATWYLEAVCPGQYGIPEPMNPFRRSYWCGPRPKSPAPCPTPLD- 745
Db 764 SVSNMFLDFTFYGVMTWYIEAVFPQYGIPEPMNPFRRSYWCGPRPKSPAPCPTPLD- 823
Qy 746 PKVLVEAPPGLSPGVSVRSLEKFPSPQALRGLSLDFYQCHITAFLGHNGAGKTTTL 805
Db 824 SEICMEEEPTHLKLGVSIGNLVKQYVRDGMKVAVDGLALNFYEGQITSEFLGHNGAGKTTM 883
Qy 806 SILSGLFPSPGGSAFILGHVDRSSMAAIRPHLGVCPQVNVFLDMLTVDEHVMWFYGRKGL 865
Db 884 SILTGLPPTSGTAYIILKDIRSEMSIRQLWGVCPQHNVFLDMLTVEEHFWFARLKL 943
Qy 866 SAAVVGPEQDRLQDVGL-VSKQSVQTRHLSSGGMQRKLSVAIAFVGGSVQVILDEPTAGV 924
Db 944 SEKHVKAEMEQLADVGLPPSKLKSQTSQSGGMQRKLSVALA FVGGSKVILDEPTAGV 1003
Qy 925 DPASRRGIWELLKYRGRITLILSTHLDREALLGDRVAVVAGGRGCCGSPFLRRHLG 984
Db 1004 DPYRRGIWELLKYRGRITLILSTHLDREALLGDRVAVVAGGRGCCGSPFLRRHLG 1063
```


Db 346 KAPLGIDSTRKDPISYDKRITTCNALLQSIUESNPLTKIAWRAAKPLVMGKILFTPDSP 405
Qy 257 FTRKMAQVNRTEELITLDRVREVMELGPRIFTFMDNSNVAMLQRL----- 306
Db 406 AVRRILQANSTFEELERLKLVAKEEVPQIWFYFDRSTQMTWIRDTLENPTVKGFLN 465
Qy 307 -QWDEGRQPRGGRDHWEALRSFLDPG-----SGGYNQWQADAHVGLVGTGRV 357
Db 466 SOLGBEGIT-----AEAMNLFHLKGPRESQADDMANFDWRDVFNTDRTLTSKY 516
Qy 358 TECLSIDLEAAPSAAALSVRALQALLAHRFWAGVVLGPESSDPTHTPDLPGRHVR 417
Db 517 LECLILDKPESYDEIQTQRALSLEENRFAGVVF-----PDMPYPTSALPETHVK 568
Qy 418 IKIRMDIDVVTNRKIRDFWDPGAADPLTDRLRYVMGGFVYLOQLVERAAVRVLSGANP 477
Db 569 YKIRMDIDVVEKTKIKORYWDSGRADPVEDFRIYWGGFAYLODMIEGQITRSQOQEV 628
Qy 478 RAGLYLOQMPYPCYVDDVFLVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRTWRA 537
Db 629 PVGIYLOQMPYPCYVDDVSMILNRCFPIFMVLAWIYSVMTVKISIVLEKELRLKXETLKN 688
Qy 538 MGLSRAVLWGLFSLCLGFLPLSALLVLVKGDLIPYSHGVVFLFLAAFAVATVQS 597
Db 689 QVGSNAVIMCTWFLDSFMSMSIFLLTIFIMHGRILHYNSPFIILFLFLAFSTATIMQC 748
Qy 598 FLLSAFFSRANLAAACGGIAYFSLYPLVYLCVAVWRDLRPAAGRVAASLLSPVAFGCGES 657
Db 749 FLLSTFFSRASLAACSGVIYTLVYLPHILCFAMQDRMTADLKMAVLSLLSPVAFGFTGY 808
Qy 658 LALLBEOQSGAQMHWGTRP-TADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQOYGP 716
Db 809 LARFEEQGLQWNSIGKSPMGDEDFSFLMSKMKMLLDAALYGLAWLYDQVFPNGYGP 868
Qy 717 EPWNPFRSRYCG-----PRPKSPACPPL-DPK-----VLVEAPGLSPG 760
Db 869 LPWYFLLQESYVLGGEGSTREERALEKTEPITEEMEDPEHPEGINDAFFERELEPLVPG 928
Qy 761 VSVRSLEKRFPGSPALRGLSLDFVQGHITAFGLHNGAGKTTLSILSGLPSPGGSF 820
Db 929 CVCKNLVKIFEPYSPVADRLNITYENCITAFGLHNGAGKTTLSILSGLPSPGTVL 988
Qy 821 ILGHVDRSMAAIRPHLGVCPOYNVLFOMLTVDEHWFYGRUKGLSAVVPGEQDRLLQD 880
Db 989 IGGKDIETSLDVRQSLGCMPOYNILFHLITVAEHILFYAQLKKGKSWEEAQLMEAMLED 1048
Qy 881 VCLVSKQSVOTRHLGGHOKLSVAIAFVGGQVILDEPTAGVDPASRGITWELLKXR 940
Db 1049 TGLHHRNEEAQDLGGHOKLSVAIAFVGGAKVILDEPTSGVDPPYRRSIWDLKLYR 1108
Qy 941 EGRTTILSTHLDLAEALLGDRVAVVAGGRGCCGSPFLRRLHLSGGYTLTLVKARLPLT 1000
Db 1109 SGRITIMSTHMDLADLGRDAIISQGLYCSGTFPLFKNCFGYFYLTVR-----KWN 1165
Qy 1001 NEKADTMEGSVYTRQEKNGSGSRVG--TP-----QLLALVQHWVFGARLVEEL 1049
Db 1166 IQSRTQCEGTCSASEGSTFRCPAHIDEITPEQVLDGDMVNMWQHHVPEAKLVECI 1225
Qy 1050 PHELVLVLPYGAHDSFATLRELDTRLAEURLTGYGISDTSLEEIFLKVVEBCAATD 1109
Db 1226 GQELIFLPLKNFKORVASFRELEDTLADLGLSGISDTPLEIEIFLKV----- 1277
Qy 1110 MDGSCGHLCTGIAGLDVTLRLK-MPQETALE-----NG-----EPAGSAPETDQSGPD 1160
Db 1278 -EDSUGPLFAGCTQOKENLRHPLWSPREKARQIPQSGNGCSXEPAPHEGQSSPE 1335
Qy 1161 AVGRVQ-GWALTROQLQALLKRFLLARRSRGLFPAQIVLPALFVGLALVFSLIYPPFGH 1219
Db 1336 ARSRLNTGAQLIIQHVALLVKRFHHTIRSHKDFLAQIVLPATFVFLALMLSLIYPPFGE 1395
Qy 1220 YPALRLSPMYGAQVFSFSDAPGDPGRARLEALLQKAG-----LEPPVQHSS 1269
Db 1396 YPALILHFWPMYQQYTTFFSLDQFGEQLAALADVLNKPGRNCLKEBWLPEYPCGNST 1455

Qy 1270 HRFSAPEVAEYAKVLASGNWTPESPSPACQSQCPGARRLLPDCPAAAGPPPPQAVTGS 1329
Db 1456 -PWKTPSVSPNITHLFQKQWTPKPSPCRSCTREKLTMLPECEGAGLPPQRIORS 1514
Qy 1330 GEVONLITGRNLSDFLVITYPRLVRQGLTKKQWNEVRYGGSPSLGGRDGLP-SQELGR 1388
Db 1515 TEILOQLTNRNISDFLVITYPALIRSSLSKFWNEQRYGGISIGGLPLIPITGEALVE 1574
Qy 1389 SVEELWALLSLPLGGALDR-VLKNLTAWAHSIDAODSLKIMFNKNGHSMVAFVNRASNA 1447
Db 1575 FLHLGQIMN-VSGGPITREASKEMPAFLKHELETDNIKWPNNGWHALVSFLVAHNA 1633
Qy 1448 ILRAHLPFGPARHAHSITTLNHLPLMLTKEQLSEALMASSVDVLVSIQVVFAMSPVPSF 1507
Db 1634 ILRTSLHKDKNPEEYGIITVISQPLNLTKQELSEITVLVTASVDVAIVCIVFAMSPVPSF 1693
Qy 1508 TLVLEERTVRAKHLQMLGGLSPTLYWLNFLWDMCNLYLPACIIVLFLAPQQAAYAP 1567
Db 1694 VLYLTQERVNRKAKHLQFVSGVSPITYWLTNFLWDMINNYAVSAALVVGIFVGFQKCAITSP 1753
Qy 1568 ANLPALLALLLLYGWSITPLMYPASFPSPSTAVVLTCTINLFIGINGSMATFVLELPS 1627
Db 1754 ENLPALIALLLMYGWAIVPMYPASFLFDVPSATVALSCANLFIGINSSAITFLELPE 1813
Qy 1628 DQ-KLQEVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROFQSPRLWEVVGK 1686
Db 1814 NNWTLRFRNMLKLLIIFPHFCLGRGLIDLALSOAVTVVARFGEHSNTPFQWDLIGK 1873
Qy 1687 NLLAMVIOGPIFLFTLLQHR---SOLLPOPRVRSPLPLGSEEDVARERERVVOGATQ 1743
Db 1874 NLVMAAEGVYVLLLTFLFIQHFFLTRWVSEPAKE--PII-DEDDVAEERQIIISGNGK 1930
Qy 1744 GDVLVLRNLTKVYQORMPAVDRLCLGIPPECECLGCVNGAGKTSTFRWVGTDLASRG 1803
Db 1931 TDILRLNELTKIYSGTSSPAVDRLCVGRPECEGLGCVNGAGKTTTFKMLTGDTVTGSG 1990
Qy 1804 EAVLAGHVAAREPAAHLSMGYCPQSDAIFELLTGREHLELIARLARGVPEAQVATAGSG 1863
Db 1991 DATIAGKILTNISDVHSGMGCYCFQDAVDLLTGREHLYLYARLARGVPADEIERNWS 2050
Qy 1864 LARGLSWADRPAGTYSGGNKRLATALALVDPVAVFLDEPTTGMPSPARRFLWNSLL 1923
Db 2051 IQSLGLSLYADRLVGTYSNGNKRKLSTAIALMGCPPLVLLDEPTTGMPQARRMLWNTIV 2110
Qy 1924 AVVREGSVMLTSHSMECEALCSRLATMNGRFRCLGSPHLLKGRFAAGHTLTLRVPA 1983
Db 2111 SIIREGRAVLTSHSMECEALCTRLAIMVKGTFQCLGTIOHLKYFGDGYIVTMKIKSP 2170
Qy 1984 RS-----QPAAAFVAAREPFGSELREAHGGRRLRFPOLPGRCALARVFGELAVHGAHGV 2037
Db 2171 KEDLLPDLNFPVEQFQGNFPGSVQVRHYNNMLQFV---SSSSLARIIFLLISHKDSLLI 2227
Qy 2038 EDFSVSQMLBEVFLYFSKQDGKDED 2063
Db 2228 BEYSVTQTLTLDQVFNFAKQQTETHD 2253
RESULT 12
AAAR87835 PRELIMINARY; PRT; 2269 AA.
AC AAR87835;
DT 20-MAY-2004 (TremBLrel. 27, Created)
DT 20-MAY-2004 (TremBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TremBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS23;


```
QY 1744 GDLVLRNLTQYRQMPAVDLCLGIPGBCFGLGVNGAGKSTFRMVTGDTLASRG 1803
D 1931 TDILALNELTKIYSTSSPAVRLCVGRPGCFGLGVNGAGKTTTFKMLTGDITVTSG 1990
QY 1804 EAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVATAGSG 1863
D 1991 DATIAGKSLTWNISVHOSMGYCPQSDAIFELLTGREHLELLARLGRVPEADIERVANWS 2050
QY 1864 LARLGLSWADRPAGTYSGNKRKLATLALVGDPAVFLDEPTTGMDDPSARRFLWNSLL 1923
D 2051 IOSLGLSLYADRLVGTYSNGNKRKLSTALMGCPPLVLLDEPTTGMDDPSARRFLWNTIV 2110
QY 1924 AVVREGSRVMTLSHMECEALCSRLAINVNGRFRCLGSPQLKGRFAAGHTLTILRVPA 1983
D 2111 SIIREGRAVLTLSHMECEALCSRLAINVNGRFRCLGSPQLKGRFAAGHTLTILRVPA 2170
QY 1984 RS-----QPAARFAVAAFPSELEBAGGRRLRFQPGRCALARVFGELAVHGAHGV 2037
D 2171 KEDLLPDLNPVQFQGNPPGVSQVREHYNMLQFV---SSSLARIIFRLTISHKDSLLI 2227
QY 2038 EDFSYSQTMLEBVFYFSKQDKD 2063
D 2228 EYSVTQTLDOQVFNFAKQQTETHD 2253
```

```
RESULT 13
Q6T940 PRELIMINARY; PRT; 2269 AA.
AC Q6T940;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN (1) SEQUENCE FROM N.A.
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.;
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.";
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY427779; AAR87836.1; --
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005951; Rim_ABC_transp.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR01257; rim_protein; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2269 AA; 256491 MW; 9E6E349FF17895A1 CRC64;
```

```
Query Match 45.0%; Score 4903; DB 2; Length 2269;
Best Local Similarity 45.5%; Pred. No. 2.3e-285;
Matches 1023; Conservative 347; Mismatches 654; Indels 222; Gaps 37;
QY 1 PPLEHHECHFNKPLPSAGTVPWLOGLICNVNNTCFPOLTPCEBGRSLNFDNLSVRL 60
D 47 PLYSQHECHFNKAMPAGMLPWLQGMFCNVNNTCFQNPPTGESPGIVSNYNNILARVF 106
```

```
QY 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPLEBPPMLDVAE----- 112
D 107 RDFQELLDDAPERQHFHGWKEFQTLRLMDTLRTHPERVAGRIRINDVLKDEETLTLF 166
QY 113 -----LLTSLLRTESI----- 123
D 167 LMKNIGLSDSVVYLLNSQVRPEQPAHGVDPDLMLKDIACSETLLRPFIFSORRGAQTVR 226
QY 124 ---GLAQQAQ-----EPLHSLLEAAEDLAQELLALRS-----LVELRA 159
D 227 DAMCSLSQGTLOWVEDTLYANVDFPKLFRVLPTLLDSSQGINLRSWGRVDFSDISRIKE 286
QY 160 LLQRP-----RGTSQP-----LELLSEALCSV-RGPSTTVGSPSLNNYEASDL 200
D 287 FIIHRPSVDELLWVTKPLTQTGPETFAQLMSILSDLLCGYPGGGSRVP-SPNMYEDNNY 345
QY 201 MELVQEPESALP-----DSSLSPACSELI GADSHPLSRLLRRLKPLILGLKLLFADPTP 256
D 346 KAFGLIDSTRKDPYSDYDKRTTTCNALIQSLESNPLTKIAWRAAKPLVMGKILFDPSP 405
QY 257 FTRKLMAQVNRTEELTLRLDRVREVMGLPRIFTFMNDSSNVAMLQRL----- 306
D 406 AVRRILQANSTFEELERLKLKVAWEVGFQIWFYFDRSTQMTWIRDTLENPTVKGFLEN 465
QY 307 -QMDEGRRQPRPGCRDHMEALRSFLDPG-----SGGYSWQDAHADVGHVGLGRV 357
D 466 SOLGEGIT-----AEAMLNFKHGPRESQADDMANFDWRDVPNITDRTLRLTSKY 516
QY 358 TECLSLDKLEAAPSEALVSALQLLAHRFWAGVVFGLGPDSSDPTBHPDLPDGHVR 417
D 517 LECLILDKFESYDDEIQLTORALSLEENRFWGVF-----PDNYPMTSALPTHVK 568
QY 418 IKIRMDIDVTRTKIRDRFWDPGPADPLDLRYVMGCVFVYLODLVRAAVRVLSGANP 477
D 569 YKIRMDIDVVKTKIKDRYWDSPRADPVDPRYINGGFAYLQDMIQGGITRSQAQVXV 628
QY 478 RAGLYLQOMPYPYCVYDDVFLRSLPLFLTLAMIYSVTLTVKAVREKTRLDTRMRA 537
D 629 PVGIYVQMPYPCFVDDSMIILNRCPPIFMVLAMIYSVMTKVSIVLEKELRLKTELKN 688
QY 538 MGLSRVNLWGLFSLCGLPFLLSAALLVLVLKGLDILPYSHPGVVFLLAFAVATVTSQ 597
D 689 QGVSNVAVICTWFLDSFMSMSIFLLTIFMHGRILHYSNPFILFLFLAFSTATIMQC 748
QY 598 FLLSAFTSRANLAACGLAYFSLVPLVLCVNRDLRPGCRVAASLLSPVAFGCGES 657
D 749 FLLSTFSRSLAAACSGVIYFTLYLPHILCFANQDRMTADLKRAVLSLSPVAFGCGTEY 808
QY 658 LALLEEGEGEAQWNVGTRP-TADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYQYGP 716
D 809 LARFEEQGLQWSNIGKSPMEGDEFSFLMSKMWLLDAAALYGLAWYLDQVFPNGYGT 868
QY 717 EPNPFPFRSYWCG-----PRPKSPACPTPL-DPK-----VLVEAPPLSGP 760
D 869 LPWYFLLQESYWLGGEGCSTREERALEKTEPTEEMEDPEHPGEGINDAFFRELGLVPG 928
QY 761 VVSRLSKRPSPQALRGLSLDFYQGHITAFIHNAGAGKTTLSILSLGFPSPGGSF 820
D 929 VCVNKLKIFEPYSRPAVDRLNITFENQIITAFIHNAGAGKTTLSILSLGFPSPGVL 988
QY 821 ILGHVRSMAAIRPHLGVCQYNVFLDMLTVDEHVMFYGRKGLSAAVAVGPEQDRLLQD 880
D 989 IGGKDIETSLDAVRSQSLGMCQYNILPHLITVAEHILFYAQLKGSWEAAQLEAMEWLED 1048
QY 881 VGLVKSQVTRHLSSGMQRKLSVAIAFVGGSGVVLDEPTAGVDPPASRRIGWELLKYR 940
D 1049 TGLHKKNEEAQDLSGGMRKLSVAIAFVGGAKVILDEPTSGVDPSRRSIRWILLKYR 1108
QY 941 EGRTLILSTHLLDRAELLGDRVAVVAGRILCCCGSPFLRRHLGSGYVYTLTVKARLPLTT 1000
D 1109 SGRITIMSTHMDADLLGDRIALIIQGRLYCSGTPLFKKNCFGTGFYTLTVR---KWN 1165
QY 1001 NEKADTMEGSDVTRQEKKGSGQSRVG--TP-----QLLALVQHVPFGARLVEEL 1049
```

Db 1166 IQSQTGCEGTCSCASGFSRCFAHIDEITPEQVLDGVNLMQVHQHVPKAVECI 1225
Qy 1050 PHELVLVLYTGAHGSFATLRELDRLAELRLTYGIGSDTSLEELFLKVVECAADTD 1109
Db 1226 QGELIFLLPNKFKORAYASLRELEDTLADGLSSFGISDTPEELFLKVT----- 1277
Qy 1110 MEDSCGQHLCTGAGLDVTLRLK-MPQETALE-----NG---EPAGSAPETDQSGPD 1160
Db 1278 -EDSDSGPLFAGGTQCKRENLRHPLSPREKARIQPGSNGSCSKEPA-PHEGQPSSEPE 1335
Qy 1161 AVGRVO-GWALTROQLQALLKRLFLARRSRGLFAQILVLPALVGLVFLSLVPPFGH 1219
Db 1336 ARSENTGAQIIHQVQALLVRRFHTIRSHKOFLOAQIVLPATVFLALMSLIIVPFGE 1395
Qy 1220 YPALRLSPTMYGAQVFFSDEAPDQGRARLEALLQEAG-----LEEPPVQHS 1269
Db 1396 YPALILHPMYQGYTFFSLDQPGSEQLAALADVLLNKPFGNRCLEKEMLPPEYPCGNST 1455
Qy 1270 HRPSAPEVPAEVAKVILASGNWTPSPACOCQPGARRLLPDCPAAAGGPPPPQAVTGS 1329
Db 1456 -PWKTPSVSPNITLHFKQQTPEKPSPCSCSTREKLTMLPECEGAGGLPPPPQRIORS 1514
Qy 1330 GEVQNLTRNLSDFLVKTYPRLVROGLKTKKWNVRYGSGSLGGRDPGLP-SGQELGR 1388
Db 1515 TEILOQDITNRNISDFLVTYPALIRSSLSKFWNEQRYGSGISGGKLPILPITGEALVE 1574
Qy 1389 SVEELWALLSPGGALDR-VLNKLTAWHSLDAQDSLKIKWFNNKGWHSVAFVNRASNA 1447
Db 1575 FLSHLGQIMN-VSGGPTREASKEMPAFLKHELTEDNIKWFNNKGWHALVSLNVAHNA 1633
Qy 1448 ILRAHLPPGPARHAHSITTLNPLNLTKQELSEALMASSVDVLVSVICVFMASFVASF 1507
Db 1634 ILRTSLHKDKNPEYGYITVISOPLNLTKQELSEITVLITASVDVAIVLCVFMASFVASF 1693
Qy 1508 TLVLIERTVTRAKHLQMLGSLPPLYLWGNPLWDMCNVLPACTVLIIFLAFOQAVAP 1567
Db 1694 VLYLIQERVNKAKHLQFVSGVSPTYLWTLNPLWDMVYSAALVGVFGFQKATSP 1753
Qy 1568 ANLPALLLLLLYGSITPLMPYASFPSFVSTAYVLTCTINLFINGSMATVLELFS 1627
Db 1754 ENLPALLALLMYGWAVIPMYPASFLFDPSTAYVALSCANLFGINSATIFILELFE 1813
Qy 1628 DQ-KLOEVSRLKOVFLIFPHFCGLRGLIDMVRNQAMADAPERLGDQFQSPLEWVVGK 1686
Db 1814 NNWTLRFNMLRKLIIFFHFCGLRGLIDLALSQAVTDVYARFGEBSHSTNPPQMDLIG 1873
Qy 1687 NLLAWIQGFLFLFTLLQHR---SOLLPOPRVRSPLPLGEEDEVDARERVRVQCATQ 1743
Db 1874 NLVMAAEGVVYLLTLFIQHFFLTRVWSEPAKE--PII-DEDDVAREQRIIISGNK 1930
Qy 1744 GDVLVRLNLTKVYRGORMPAVDRICLGIIPGECFGLLVGAGKGTSTFRMTGDTLASRG 1803
Db 1931 TDILRLNELTKIYSGTSSPAVDRLCVGRPGECFGLLVGAGKGTTFTKMLTGDTTTS 1990
Qy 1804 EAVLAGHVAAREPAAHLSNGYCYQSDAIPELLTGREHLELLAHLRGVPAQVATAGSG 1863
Db 1991 DATIAGKSILTNISDVHQSMGYCQFDVADVLLTGREHLYARLGRVPADEITERVANWS 2050
Qy 1864 LARLGLSWYADRPAGTYSGGNKRKLATALALVGPVAVVFLDEPTTGMDDPSARRFLWNSLL 1923
Db 2051 IQSLGLSYADRVLVTYSGGNKRKLSTALAMGCPVLLDEPTTGMDDPSARRFLWNTIV 2110
Qy 1924 AVVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAA 1983
Db 2111 SIIREGRAVLTSHSMECEALCRLAIMVNGTFCQCLGTIQLHYKFGDGYIVTMKIKSP 2170
Qy 1984 RS-----OPAAAFVAEFPQSELRBAHGGHRLRQLPPGRCALARVFGELAVHGAHGV 2037
Db 2171 KEDLLPOLNVEQFPQGNFPGSVQORRHYNMLQVQ---SSSSLIARIFRLLISHKDSLLI 2227
Qy 2038 EDFSVSOTMLREVFYFSKQGDDED 2063

Db 2228 EEYSVTQTTLDQVFNFAKQQTETHD 2253
RESULT 14
AAR87836
ID AAR87836 PRELIMINARY; PRT; 2269 AA.
AC AAR87836;
DT 20-MAY-2004 (TREMREL. 27, Created)
DT 20-MAY-2004 (TREMREL. 27, Last sequence update)
DT 20-MAY-2004 (TREMREL. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.,
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.",
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.,
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY427779; AAR87836.1; -;
SQ SEQUENCE 2269 AA; 256491 MW; 986E349FF17895A1 CRC64;
Query Match 45.0%; Score 4903; DB 2; Length 2269;
Best Local Similarity 45.5%; Pred. No. 2.3e-285;
Matches 1023; Conservative 347; Mismatches 654; Indels 222; Gaps 37;
Qy 1 PPLEHHECHPNKPLPSAGTVPMLOGLICNNVNTCPQLTPGEPGLSNDLSVRL 60
Db 47 PLYSQHECHPNKAMPAGMLPWLQGMFCNNVPCQNPPTGSPGIVSYNNILARV 106
Qy 61 ADARTVLGASARTLAGLKLATLRAARSTAQPOPTKOSPPLPMLDVAE----- 112
Db 107 RDQELLLDAPERQHGHVKEQTLRLMDLRTHPVAGRGIRRDVLKDEETLFL 166
Qy 113 -----LTSLLRTESL----- 123
Db 167 LMKNIGLSDSVWYLLNSQVRPEQFAHGVDPDMLKDIACSETLLERFIISQRRGAQTVR 226
Qy 124 ----GLALGQAO-----EPLHSLLEAEGLAQELLARS-----LVELRA 159
Db 227 DAMCSLSQGTLOWVEDTLANVDFKLFRLVLTLLSSSQGINLRSGRVFSOISSIRE 286
Qy 160 LLORP-----RGTSGP-----LELLSEALCSV-RGPSSTVGPSLNWEASDL 200
Db 287 FIHRPSVEDLLWTKPLTQTGGPETAQLMSILSDLLCGYPEGGSRVF-SFNWYEDNY 345
Qy 201 MELVQGPESALP-----DSSLSPACSELIGALDSHPLSRLMLRRLKPLILGLKLLFADPTP 256
Db 346 KAFGLIDSTRKDPITYSDKRTTTFNCALIQSLSENPLTKIARAAPKPLVMGKILFTPDSP 405
Qy 257 FTRKLAQVNRTEELTLDRDREVWMLGPRTFTMNDSSNVMLQRL----- 306
Db 406 AVRIIQNANSTFEELERLKLKVAWEVGPQIWFDFRSTQMTIMRTDLENPTVKGFLN 465
Qy 307 -QMODERRQPRPGRDHMEALRSFLDPG-----SGGYSQDAHADVHGLVGLTGRV 357
Db 466 SQEGEGIT-----AEAMLNFLHKGPRESQADMANFDRVFNIDRTLRLTSKY 516
Qy 358 TECLSDKLKFAAEEAALVSRLAQLAEHFWAGVGVFLGPDSDSDTEHPTDLPGRHVR 417
Db 517 LECLILDKFESYDDEIQLTORALLSLEENRFWAGVVF-----PDMYPWTSALPETHVK 568

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:28:54 ; Search time 106.335 Seconds
(without alignments)
7104.204 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHHECHFPNKLPLSAGT.....QHKRVSQFLDDPSTAETVL 2100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10896	100.0	2100	9 US-09-995-542-6	Sequence 6, Appli
2	10896	100.0	2146	9 US-09-995-542-5	Sequence 5, Appli
3	10883	99.9	2146	15 US-10-182-006-2	Sequence 2, Appli
4	10883	99.9	2180	15 US-10-332-447-27	Sequence 27, Appli
5	10880	99.9	2144	9 US-09-858-194-2	Sequence 2, Appli
6	10880	99.9	2144	14 US-10-154-419-2	Sequence 2, Appli
7	10874.5	99.8	2147	17 US-10-618-281-44	Sequence 44, Appli
8	10392.5	95.4	2059	15 US-10-114-270-176	Sequence 176, Appl
9	9213.5	84.6	1873	15 US-10-182-006-4	Sequence 4, Appli
10	8389.5	77.0	2121	9 US-09-995-542-3	Sequence 3, Appli
11	8389.5	77.0	2167	9 US-09-995-542-2	Sequence 2, Appli
12	7596	69.7	1550	9 US-09-995-542-8	Sequence 8, Appli
13	5610.5	51.5	2261	14 US-10-313-641-9	Sequence 9, Appli

14	5610.5	51.5	2261	14	US-10-313-641-10	Sequence 10, Appli
15	5610.5	51.5	2261	14	US-10-428-551-9	Sequence 9, Appli
16	5610.5	51.5	2261	14	US-10-428-551-10	Sequence 10, Appli
17	5610.5	51.5	2261	15	US-10-452-510-1	Sequence 1, Appli
18	5610.5	51.5	2261	15	US-10-617-334-1	Sequence 1, Appli
19	5610.5	51.5	2261	16	US-10-745-377-5	Sequence 5, Appli
20	5610.5	51.5	2261	16	US-10-744-465-1	Sequence 1, Appli
21	5610.5	51.5	2261	16	US-10-833-679-1	Sequence 1, Appli
22	5610.5	51.5	2261	17	US-10-872-113-5	Sequence 5, Appli
23	5603.5	51.4	2261	14	US-10-340-097-118	Sequence 118, App
24	5603.5	51.4	2261	14	US-10-336-215-118	Sequence 118, App
25	5603.5	51.4	2261	14	US-10-336-219-118	Sequence 118, App
26	5594.5	51.3	2261	10	US-09-984-827-128	Sequence 128, App
27	5594.5	51.3	2261	10	US-09-984-827-132	Sequence 132, App
28	5594.5	51.3	2261	10	US-09-984-827-134	Sequence 134, App
29	5594.5	51.3	2261	10	US-09-984-827-136	Sequence 136, App
30	5593.5	51.3	2261	9	US-09-995-542-11	Sequence 11, Appli
31	5593.5	51.3	2261	9	US-09-846-456-11	Sequence 2, Appli
32	5593.5	51.3	2261	10	US-09-984-827-2	Sequence 127, App
33	5593.5	51.3	2261	10	US-09-984-827-127	Sequence 129, App
34	5593.5	51.3	2261	10	US-09-984-827-129	Sequence 130, App
35	5593.5	51.3	2261	10	US-09-984-827-130	Sequence 131, App
36	5593.5	51.3	2261	10	US-09-984-827-131	Sequence 133, App
37	5593.5	51.3	2261	10	US-09-984-827-133	Sequence 2326, Ap
38	5588.5	51.3	2261	15	US-10-276-774-2326	Sequence 9, Appli
39	5586.5	51.3	2263	15	US-10-276-774-2326	Sequence 293, App
40	5531.5	50.8	2201	9	US-09-995-542-9	Sequence 29, Appli
41	5522.5	50.7	2201	14	US-10-170-385-293	Sequence 4, Appli
42	5522.5	50.7	2201	14	US-10-331-496A-29	Sequence 139, App
43	5522.5	50.7	2201	15	US-10-429-160-4	Sequence 140, App
44	5522.5	50.7	2201	16	US-10-465-498-139	
45	5470.5	50.2	2233	16	US-10-465-498-140	

ALIGNMENTS

RESULT 1

US-09-995-542-6
; Sequence 6, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-6

Query Match 100.0%; Score 10896; DB 9; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PPLEHHECHFPNKLPSAGTVPWLOGLICVNNVTCFFPOLTGEPRGLSNFNDLSVRL 60
Db	1	PPLEHHECHFPNKLPSAGTVPWLOGLICVNNVTCFFPOLTGEPRGLSNFNDLSVRL 60
Qy	61	ADARTVLGGASAHRTLAGLGLIATRAARSTAOPTKQSPLEPPMDVAELTSLRT 120
Db	61	ADARTVLGGASAHRTLAGLGLIATRAARSTAOPTKQSPLEPPMDVAELTSLRT 120
Qy	121	ESLGALGQAEPLHSLLEAEDLAQELLARLSIVELRALQRPRTGSGPLELLEALCS 180
Db	121	ESLGALGQAEPLHSLLEAEDLAQELLARLSIVELRALQRPRTGSGPLELLEALCS 180

Db 121 ESGLGALGOAQBPLHSLLEAAEDLAQELLALRSIVELRALLQRPGTSGPLELLSEALCS 180
Qy 181 VRGFSSTVGPSSLNWEASD.MEILVGPBPESALPDSSLPACSEBILGALDSHPISRLWRR 240
Db 181 VRGFSSTVGPSSLNWEASD.MEILVGPBPESALPDSSLPACSEBILGALDSHPISRLWRR 240
Qy 241 LKPLILGLKLLFADPTPTPKLMAQVNRTPFELTLRLDRVREWMGPRIFTFMNDSSNVA 300
Db 241 LKPLILGLKLLFADPTPTPKLMAQVNRTPFELTLRLDRVREWMGPRIFTFMNDSSNVA 300
Qy 301 MLQRLQOMQDEGRQRPGRDHEALRSFLDPGSGYSQDADADVGLHVLGTLGRVTEC 360
Db 301 MLQRLQOMQDEGRQRPGRDHEALRSFLDPGSGYSQDADADVGLHVLGTLGRVTEC 360
Qy 361 LSLDKLEAPSEALVSRAQLLAHRFWAGVVFGLGEDSSDPTEHPTPDLGPHGVRIKI 420
Db 361 LSLDKLEAPSEALVSRAQLLAHRFWAGVVFGLGEDSSDPTEHPTPDLGPHGVRIKI 420
Qy 421 RMDIDVTRINKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 421 RMDIDVTRINKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
Qy 481 LYLQOMPYPCYVDDVFLRVLRSPLFLTLAWIYSVTLTVKAUVREKETRLRDTMRAMGL 540
Db 481 LYLQOMPYPCYVDDVFLRVLRSPLFLTLAWIYSVTLTVKAUVREKETRLRDTMRAMGL 540
Qy 541 SRAVLWLGWFLSCLGPFLLSAALLVLVLKLGDLIPYSHPGVWFLFAAFVAVATVTSFLL 600
Db 541 SRAVLWLGWFLSCLGPFLLSAALLVLVLKLGDLIPYSHPGVWFLFAAFVAVATVTSFLL 600
Qy 601 SAFFSRANLAAACGLAYFSLYLPPYVLCVAVRDRLPAGGRVAASLLSPVAFGCGCSLAL 660
Db 601 SAFFSRANLAAACGLAYFSLYLPPYVLCVAVRDRLPAGGRVAASLLSPVAFGCGCSLAL 660
Qy 661 LEQGEQAQHNHVTGRTADVFSLAQVSGLLDAAALYGLATWYLEAVCPGQYGIPEPN 720
Db 661 LEQGEQAQHNHVTGRTADVFSLAQVSGLLDAAALYGLATWYLEAVCPGQYGIPEPN 720
Qy 721 FPFRRSYWCGPRPPKSPACPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 721 FPFRRSYWCGPRPPKSPACPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
Qy 781 LSLDFYQCHITAFILGHNGAGKTTLSLSGLFPSPGSAFILGHVDRSSMAAARPHLGV 840
Db 781 LSLDFYQCHITAFILGHNGAGKTTLSLSGLFPSPGSAFILGHVDRSSMAAARPHLGV 840
Qy 841 PQYNVLFDMLTVDBHWFYGRKLGLSAAVVGPEODRLLODVLGYSKQSVOTRHLSGQMOR 900
Db 841 PQYNVLFDMLTVDBHWFYGRKLGLSAAVVGPEODRLLODVLGYSKQSVOTRHLSGQMOR 900
Qy 901 KLSVAIAFVGSQVVIIDEPAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
Db 901 KLSVAIAFVGSQVVIIDEPAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
Qy 961 RVAVVAGGRICCCGSPFLRRHLGSGYLLVLVKARLPLTTNEKADTDMEGSVDTROEKN 1020
Db 961 RVAVVAGGRICCCGSPFLRRHLGSGYLLVLVKARLPLTTNEKADTDMEGSVDTROEKN 1020
Qy 1021 GSQSSRVGTTPQLLALVOHVPGARLVEPHELVLVLPYTGADHGSFATLIFRELDTRLAE 1080
Db 1021 GSQSSRVGTTPQLLALVOHVPGARLVEPHELVLVLPYTGADHGSFATLIFRELDTRLAE 1080
Qy 1081 LRLTYGIDSLSLEEIFLKVVEECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPQETA 1140
Db 1081 LRLTYGIDSLSLEEIFLKVVEECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPQETA 1140
Qy 1141 LENGEPAAGSAPETQCGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRGILFAQIVLP 1200
Db 1141 LENGEPAAGSAPETQCGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRGILFAQIVLP 1200
Qy 1201 ALFVGLALVLSLIVPPFGHPALRLSPMTYGAQVSFFSEADPGPGRARLLEALLQAGL 1260
Db 1201 ALFVGLALVLSLIVPPFGHPALRLSPMTYGAQVSFFSEADPGPGRARLLEALLQAGL 1260

Qy 1261 EBPVQVSHRRSAPVPAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1320
Db 1261 EBPVQVSHRRSAPVPAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1320
Qy 1321 PPOAVTGSGEVVQNLTGRNLSDFLVKTYPRLVQRQLTKKWVNEVRYGGFSLGGRDPGL 1380
Db 1321 PPOAVTGSGEVVQNLTGRNLSDFLVKTYPRLVQRQLTKKWVNEVRYGGFSLGGRDPGL 1380
Qy 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIWFNNKQWHSMVAF 1440
Db 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIWFNNKQWHSMVAF 1440
Qy 1441 VNRASNAIIRAHLPPOPARHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAM 1500
Db 1441 VNRASNAIIRAHLPPOPARHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAM 1500
Qy 1501 SFVPASFTLVLIERVTRAKHLQMGGLSFTLYWLNFLNWMCNLYLVPACIVVLIFLAFQ 1560
Db 1501 SFVPASFTLVLIERVTRAKHLQMGGLSFTLYWLNFLNWMCNLYLVPACIVVLIFLAFQ 1560
Qy 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPAGFFSVPTAYVVLTCINLFIGINGSMAT 1620
Db 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPAGFFSVPTAYVVLTCINLFIGINGSMAT 1620
Qy 1621 FVLEIFSDQKLQVRSILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDFOFOSPLR 1680
Db 1621 FVLEIFSDQKLQVRSILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDFOFOSPLR 1680
Qy 1681 WEVVGKNLLAMVTOGFLFLLFTLLQHRSQLPQPRVRSILPLIGEEDVAVRERVRVQG 1740
Db 1681 WEVVGKNLLAMVTOGFLFLLFTLLQHRSQLPQPRVRSILPLIGEEDVAVRERVRVQG 1740
Qy 1741 ATQGDVVLRLNTKVYRGORMPAVDRLCLGIPGECFGLLVGNVAGAKTSTFRMVTGDTLA 1800
Db 1741 ATQGDVVLRLNTKVYRGORMPAVDRLCLGIPGECFGLLVGNVAGAKTSTFRMVTGDTLA 1800
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPSQDAIPELLTGREHLELLARLRGVPEAQVOTA 1860
Db 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPSQDAIPELLTGREHLELLARLRGVPEAQVOTA 1860
Qy 1861 GSGLARLGLSWYADRPAQYVSGNKKELATALAVGDPVAVFVDEPTTGMDDPSARRFLWN 1920
Db 1861 GSGLARLGLSWYADRPAQYVSGNKKELATALAVGDPVAVFVDEPTTGMDDPSARRFLWN 1920
Qy 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Db 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Qy 1981 PAARSQPAAFVAAEFPGSELREAHGRLRFQIPGGRCALARVFGELAVHGAHGVEDF 2040
Db 1981 PAARSQPAAFVAAEFPGSELREAHGRLRFQIPGGRCALARVFGELAVHGAHGVEDF 2040
Qy 2041 SVSOTMLEEVFLYFSKQDKDETEBQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAEVTL 2100
Db 2041 SVSOTMLEEVFLYFSKQDKDETEBQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAEVTL 2100

RESULT 2

US-09-995-542-5

; Sequence 5, Application US/09995542

; Patent No. US20020127647A1

; GENERAL INFORMATION:

; APPLICANT: Shutter, John

; APPLICANT: Ulias, Laarni

; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and

; TITLE OF INVENTION: Uses thereof

; FILE REFERENCE: 00-658-A

; CURRENT APPLICATION NUMBER: US/09/995,542

; PRIOR FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: 60/253,520

; PRIOR FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-5

Query Match      100.0%; Score 10896; DB 9; Length 2146;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRISNFNDSIVSRLL 60
DB 47 PPLEHHECHFPNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRISNFNDSIVSRLL 106
QY 61 ADARTVLGASAHRTLAGLGLKIATLRAARSTPAQOPTKQSPLEPMPDVAELTSLT 120
DB 107 ADARTVLGASAHRTLAGLGLKIATLRAARSTPAQOPTKQSPLEPMPDVAELTSLT 166
QY 121 ESLGIALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSGLLELSEALCS 180
DB 167 ESLGIALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSGLLELSEALCS 226
QY 181 VRGSPSTVGPNSLWTEASDLMELVQCEPESALPDSSLSAPACESLIGALDSDHPLSKLLWRR 240
DB 227 VRGSPSTVGPNSLWTEASDLMELVQCEPESALPDSSLSAPACESLIGALDSDHPLSKLLWRR 286
QY 241 LKPLTLGLKLPADPTFTTKLMAQVNRFTFEELTLARDVREVWMLGPRIPTFMNDSSNVA 300
DB 287 LKPLTLGLKLPADPTFTTKLMAQVNRFTFEELTLARDVREVWMLGPRIPTFMNDSSNVA 346
QY 301 MLQRLLOQDQGRQPRPGGRDHMEALRSFLDPGSGGYSWODAHADVGHVLTGLRVTEC 360
DB 347 MLQRLLOQDQGRQPRPGGRDHMEALRSFLDPGSGGYSWODAHADVGHVLTGLRVTEC 406
QY 361 LSLDKLEAAPSEAAALVSRALQALLABHRFWAGVFLGPDSDSDPTBHTPDIGPGHVRKI 420
DB 407 LSLDKLEAAPSEAAALVSRALQALLABHRFWAGVFLGPDSDSDPTBHTPDIGPGHVRKI 466
QY 421 RMDIDVTRTKIRDFWDGPAADPLTDLRVVWGFFVYLODLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTKIRDFWDGPAADPLTDLRVVWGFFVYLODLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCYVDDVFLVLSRSLPLTLTAWISVTLTVKAVREKETRLDRTRAMGL 540
DB 527 LYLOQMPYPCYVDDVFLVLSRSLPLTLTAWISVTLTVKAVREKETRLDRTRAMGL 586
QY 541 SRAVLWLGWFLSCLGPFLLSALLVLVLKGLDILPYSHPGVVVFLFLAAFAVATVTSFLL 600
DB 587 SRAVLWLGWFLSCLGPFLLSALLVLVLKGLDILPYSHPGVVVFLFLAAFAVATVTSFLL 646
QY 601 SAFFSRANLAAACGLAYFSLYLPVLCVARNDRLPAGGRVAASLLSPVAFGFCESLAL 660
DB 647 SAFFSRANLAAACGLAYFSLYLPVLCVARNDRLPAGGRVAASLLSPVAFGFCESLAL 706
QY 661 LEEQEGEAQHNWVTRPTADVFLSAQVSGLLLDLAALYGLATWYLEAVCPGOYGIPEPWN 720
DB 707 LEEQEGEAQHNWVTRPTADVFLSAQVSGLLLDLAALYGLATWYLEAVCPGOYGIPEPWN 766
QY 721 FPFRRSYWCGPRPPKSPAPCTPLDPKVLVEEAPFGLSPGVSVRSLEKRFPGSPQPALRG 780
DB 767 FPFRRSYWCGPRPPKSPAPCTPLDPKVLVEEAPFGLSPGVSVRSLEKRFPGSPQPALRG 826
QY 781 LSLDPYQGHITAFIAGHNGAGKTTTILSGLPPSPGGSFAFILGHDRVSRSSMAAIRPHLQVC 840
DB 827 LSLDPYQGHITAFIAGHNGAGKTTTILSGLPPSPGGSFAFILGHDRVSRSSMAAIRPHLQVC 886
QY 841 POYNVLFQMLTVDEHWFVGRKGLSAAVQGEQDRLLQDVLGSKVQSVQTRHLSGGMQR 900
DB 887 POYNVLFQMLTVDEHWFVGRKGLSAAVQGEQDRLLQDVLGSKVQSVQTRHLSGGMQR 946
QY 901 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
DB 947 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 1006
QY 961 RVAVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1020
DB 1007 RVAVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1066
QY 1021 GSQSRVGTPLLALVQHWVPGARLVEBELVVLVLPYTGADHGSFATLFRLELTRLAE 1080
DB 1067 GSQSRVGTPLLALVQHWVPGARLVEBELVVLVLPYTGADHGSFATLFRLELTRLAE 1126
QY 1081 LRLTGYGISTDSLSBEIFLKVVVECAADTMDGSGCGHLCCTGIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGYGISTDSLSBEIFLKVVVECAADTMDGSGCGHLCCTGIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAASAPETDQSGPDVAVRQGWALTRQOLQALLKRFLLARRSRRLGFLAQIIVLP 1200
DB 1187 LENGEPAASAPETDQSGPDVAVRQGWALTRQOLQALLKRFLLARRSRRLGFLAQIIVLP 1246
QY 1201 ALFVGLALVFSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDPAGDPGRARLLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDPAGDPGRARLLLEALLQEAGL 1306
QY 1261 BEPPVQSHSRFSAPEVAEVAKVLASGNWTPSPSPACQSQPCARLLPDCPAAAGP 1320
DB 1307 BEPPVQSHSRFSAPEVAEVAKVLASGNWTPSPSPACQSQPCARLLPDCPAAAGP 1366
QY 1321 PPOAVTGSGBVQNLGTGRNLSDFLVKTYPRLVQGLTKKWNVEVRVYGGFSLGGRDGL 1380
DB 1367 PPOAVTGSGBVQNLGTGRNLSDFLVKTYPRLVQGLTKKWNVEVRVYGGFSLGGRDGL 1426
QY 1381 PSGQLGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIWFNKNKWSMVA 1440
DB 1427 PSGQLGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIWFNKNKWSMVA 1486
QY 1441 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSVVAM 1500
DB 1487 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSVVAM 1546
QY 1501 SFVPASFTVLIEERVTRAKHLQMLGGLSPTLYWLNFLWDMCNVLYVPACIVLLIFLAFO 1560
DB 1547 SFVPASFTVLIEERVTRAKHLQMLGGLSPTLYWLNFLWDMCNVLYVPACIVLLIFLAFO 1606
QY 1561 QRAVAPANLPAULLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT 1620
DB 1607 QRAVAPANLPAULLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT 1666
QY 1621 FVLELFSQKLQVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROQSPILR 1680
DB 1667 FVLELFSQKLQVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROQSPILR 1726
QY 1681 WEVVGKNLAMVIOGPIFLFTLLQHRSQLLPORVRSPLPLGDEEDVARERERVQ 1740
DB 1727 WEVVGKNLAMVIOGPIFLFTLLQHRSQLLPORVRSPLPLGDEEDVARERERVQ 1786
QY 1741 ATQGDVLVRLNLTKVYRGQMPAVDRCLGIPPGCEFCGLLVNGVAGKSTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLVRLNLTKVYRGQMPAVDRCLGIPPGCEFCGLLVNGVAGKSTSTFRMVTGDTLA 1846
QY 1801 SRGSAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLURGVPEAOQTA 1860
DB 1847 SRGSAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLURGVPEAOQTA 1906
QY 1861 GSGLARILGSWYADRPAGTYSGGNKRKIATALVGDPAVVPDEPTTGMPSARRFLWN 1920
DB 1907 GSGLARILGSWYADRPAGTYSGGNKRKIATALVGDPAVVPDEPTTGMPSARRFLWN 1966
QY 1921 SLLAVREGSRVMTLTSHEMCEALCSRLAIVNNGRFRCLGSPQHLKGRFAAGHTLTVR 1980
DB 1967 SLLAVREGSRVMTLTSHEMCEALCSRLAIVNNGRFRCLGSPQHLKGRFAAGHTLTVR 2026
QY 1981 PAARSQPAAPFAVAEFPGSELRHAGGRRLRQLPPGRCALARVFGELAVHGAEGHVEDF 2040
DB 2027 PAARSQPAAPFAVAEFPGSELRHAGGRRLRQLPPGRCALARVFGELAVHGAEGHVEDF 2086
```

Qy 2041 SVSOTMLEEFLYESKDGKDEDETEQKEAGVGVDPAPGLOHPKRVSQFLDDSTAETVL 2100
 Db 2087 SVSOTMLEEFLYESKDGKDEDETEQKEAGVGVDPAPGLOHPKRVSQFLDDSTAETVL 2146

RESULT 3
 US-10-182-006-2
 ; Sequence 2, Application US/10182006
 ; Publication No. US20040048250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warner-Lambert Company
 ; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
 ; TITLE OF INVENTION: THEREFROM
 ; FILE REFERENCE: National Filing
 ; CURRENT APPLICATION NUMBER: US/10/182,006
 ; CURRENT FILING DATE: 2002-07-23
 ; PRIOR APPLICATION NUMBER: PCT/US01/02191
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/177,889
 ; PRIOR FILING DATE: 2000-01-24
 ; PRIOR APPLICATION NUMBER: 60/215,405
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2146
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-182-006-2

Query Match 99.9%; Score 10883; DB 15; Length 2146;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2098; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPLEHHCCHPNKPLPSAGTVPMLOGLICNVNNTCFPOLTPGEPGRLSNFNDSLSVRL 60
 Db 47 PPLEHHCCHPNKPLPSAGTVPMLOGLICNVNNTCFPOLTPGEPGRLSNFNDSLSVRL 106

Qy 61 ADARTVLGGASAHRTLAGLKIATLRAARSTAOPTKOSPPLPMDVAELLSLRT 120
 Db 107 ADARTVLGGASAHRTLAGLKIATLRAARSTAOPTKOSPPLPMDVAELLSLRT 166

Qy 121 ESIGLALGOEPLSHLLEAEADLAQELLALRSVELRALLORPGTSGPLELLSEALCS 180
 Db 167 ESIGLALGOEPLSHLLEAEADLAQELLALRSVELRALLORPGTSGPLELLSEALCS 226

Qy 181 VRGPGSTVGPLNWEASDLMEVLGQEPESALPDSSLSPACSELIGALDGHPLSRLLWRR 240
 Db 227 VRGPGSTVGPLNWEASDLMEVLGQEPESALPDSSLSPACSELIGALDGHPLSRLLWRR 286

Qy 241 LKPLILGKLLFAPDTPTRKLMAQVNTFBEELTLRLDRVREWEMLGPRIFTFMDNSNVA 300
 Db 287 LKPLILGKLLFAPDTPTRKLMAQVNTFBEELTLRLDRVREWEMLGPRIFTFMDNSNVA 346

Qy 301 MLQRLLOQDEGRQPPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
 Db 347 MLQRLLOQDEGRQPPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406

Qy 361 LSLDKLEAAPSEALVSRALQLLAHFRFWAGVFLGPDSSDPTHEHTPDLGPCHVRIKI 420
 Db 407 LSLDKLEAAPSEALVSRALQLLAHFRFWAGVFLGPDSSDPTHEHTPDLGPCHVRIKI 466

Qy 421 RMDIDVVRTNKIRDRFWDPGPAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
 Db 467 RMDIDVVRTNKIRDRFWDPGPAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526

Qy 481 LYLOQMPCYVDVDFLRLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRMANGL 540
 Db 527 LYLOQMPCYVDVDFLRLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRMANGL 586

Qy 541 SRAVLWGLFSLCLGPFLLSAALLVLVLKGLDILPYGSHPGVVFLLAFAFVATVTSQFL 600

Db 587 SRAVLWGLFSLCLGPFLLSAALLVLVLKGLDILPCSHPGVVFLLAFAFVATVTSQFL 646
 Qy 601 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVRRDLFAGGRVAASLLSPVAFGGCESLAL 660
 Db 647 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVRRDLFAGGRVAASLLSPVAFGGCESLAL 706

Qy 661 LEEQSGGAQWHNVGTRPTADVFSLAQVSGLLDAAALYGLATWYLEAVCPGOVGIPEPWN 720
 Db 707 LEEQSGGAQWHNVGTRPTADVFSLAQVSGLLDAAALYGLATWYLEAVCPGOVGIPEPWN 766

Qy 721 FPFRRSYWCGPRPKSPAPCPTPLDPKLVVEEAPPGSLPGSVRSLEKRFPGSPQPALRG 780
 Db 767 FPFRRSYWCGPRPKSPAPCPTPLDPKLVVEEAPPGSLPGSVRSLEKRFPGSPQPALRG 826

Qy 781 LSLDFYQGHITAFIGHNGAGKTTTILSGLFPSPGSSAFILCHDVRRSSMAALRPHLVCVC 840
 Db 827 LSLDFYQGHITAFIGHNGAGKTTTILSGLFPSPGSSAFILCHDVRRSSMAALRPHLVCVC 886

Qy 841 PQYNVLFDMLTVDHEVWFYGRGLKGLSAAVVGPPQDRLLQDVGVLVSKQSVOTRHLSGQMOR 900
 Db 887 PQYNVLFDMLTVDHEVWFYGRGLKGLSAAVVGPPQDRLLQDVGVLVSKQSVOTRHLSGQMOR 946

Qy 901 KLSVAIAFVGGSQVWILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLLDEAELLGD 960
 Db 947 KLSVAIAFVGGSQVWILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLLDEAELLGD 1006

Qy 961 RVAWAGGRLCCCGSPFLARRHLSGYLTLVKARLPLITNEKADTDMESVDTROBKN 1020
 Db 1007 RVAWAGGRLCCCGSPFLARRHLSGYLTLVKARLPLITNEKADTDMESVDTROBKN 1066

Qy 1021 GSGSRVGTPLLALVQHWVPGARLVEELPHELVLPYTGADHGSFATLRELDTRLAE 1080
 Db 1067 GSGSRVGTPLLALVQHWVPGARLVEELPHELVLPYTGADHGSFATLRELDTRLAE 1126

Qy 1081 LRLTGIGISDTSLEEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1140
 Db 1127 LRLTGIGISDTSLEEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1186

Qy 1141 LENGEPAGSAPETDOGSGDDAVGVQGWALTROOLQALLKRLFLARRSRGLFAQIVLP 1200
 Db 1187 LENGEPAGSAPETDOGSGDDAVGVQGWALTROOLQALLKRLFLARRSRGLFAQIVLP 1246

Qy 1201 ALFVGLALVFLSLVPPFGHYPALRLSPTMYGAOVSFSEDPADPGARLLEALLOEAGL 1260
 Db 1247 ALFVGLALVFLSLVPPFGHYPALRLSPTMYGAOVSFSEDPADPGARLLEALLOEAGL 1306

Qy 1261 EEPVQVHSHRFSAPAEVPAEVAKVLASGNWTPESPACQSPGARRLLPDCPAAAGGP 1320
 Db 1307 EEPVQVHSHRFSAPAEVPAEVAKVLASGNWTPESPACQSPGARRLLPDCPAAAGGP 1366

Qy 1321 PPQAVTGSGEVQNLITGRNLSDFLVKTYPRLYROGLKTKWVNEVRYGFSLGGRDPGL 1380
 Db 1367 PPQAVTGSGEVQNLITGRNLSDFLVKTYPRLYROGLKTKWVNEVRYGFSLGGRDPGL 1426

Qy 1381 PSQGEIGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSIDAQDSLKTFWKNKWHSNVAF 1440
 Db 1427 PSQGEIGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSIDAQDSLKTFWKNKWHSNVAF 1486

Qy 1441 VNRASNAILRAHLPPOGARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLSICVVFAM 1500
 Db 1487 VNRASNAILRAHLPPOGARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLSICVVFAM 1546

Qy 1501 SFVPASFTVLVIERVTRAKHLOIMGGLSPTLWLGFLWDMCNLYLPACIVVLI FLAFQ 1560
 Db 1547 SFVPASFTVLVIERVTRAKHLOIMGGLSPTLWLGFLWDMCNLYLPACIVVLI FLAFQ 1606

Qy 1561 QRAYVAPANLPALLLLLLYVWSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMAT 1620
 Db 1607 QRAYVAPANLPALLLLLLYVWSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMAT 1666

Qy 1621 FVLELFSQDKLQBSRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLR 1680
 Db 1667 FVLELFSQDKLQBSRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLR 1726

QY 1681 WEVVGKLLAMVIOGFLPLFTLLQHRSQLPQVRVSLPLLGREDEDAVERERVVOG 1740
DB 1727 WEVVGKLLAMVIOGFLPLFTLLQHRSQLPQVRVSLPLLGREDEDAVERERVVOG 1786
QY 1741 ATQGDVLRNLTKVYRGQMPA VDRCLCIGIPPGECFGLGVLGNGAGKTSTFRMVTDTLA 1800
DB 1787 ATQGDVLRNLTKVYRGQMPA VDRCLCIGIPPGECFGLGVLGNGAGKTSTFRMVTDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1906
QY 1861 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARRFLWN 1920
DB 1907 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARRFLWN 1966
QY 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIVMNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
DB 1967 SLLAVVREGRSVMLTSHSMECEALCSRLAIVMNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQAAAFAAEPFSGSELBEAHGGRRLRFOLPGGRCALARVFGELAVHGAEGVEDF 2040
DB 2027 PAARSQAAAFAAEPFSGSELBEAHGGRRLRFOLPGGRCALARVFGELAVHGAEGVEDF 2086
QY 2041 SVSQTMLBEVFLYFSKQDKDBTEBEQKEAGVGVDPAPGLQHPKRVSOFLDDPSTAEVTL 2100
DB 2087 SVSQTMLBEVFLYFSKQDKDBTEBEQKEAGVGVDPAPGLQHPKRVSOFLDDPSTAEVTL 2146

RESULT 4

US-10-332-447-27

; Sequence 27, Application US/10332447.

; Publication No. US20040053258A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; RAUMANN, Brigitte B.;

; APPLICANT: THORNTON, Michael; DING, Li; YUE, Henry;

; APPLICANT: TANG, Y. Tom; HARLAND, Lee; BUREFORD, Neil;

; APPLICANT: GREENE, Barrie D.; SANJANWALA, Madhu S.;

; APPLICANT: BAUGHN, Mariah R.; YAO, Monique G.; YANG, Junming;

; APPLICANT: ARVIZU, Chandra S.; GANDHI, Ameena R.;

; APPLICANT: HAPALIA, April J.A.; TRYBOULEY, Catherine M.;

; APPLICANT: WALSH, Roderick T.; RAMKUMAR, Jayalaxmi;

; APPLICANT: LU, Yan; LU, Dyang Aina M.; AZIMZAI, Yalda;

; APPLICANT: LAU, Preeti; ELLIOTT, Vicki S.; NGUYEN, Daniel B.;

; APPLICANT: KHAN, Farrah A.; KEARNEY, Liam; THANGAVELU, Kavitha;

; APPLICANT: DAS, Debopriya; POLICKY, Jennifer L.

; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS

; FILE REFERENCE: PI-0149 USN

; CURRENT APPLICATION NUMBER: US/10/332,447

; CURRENT FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: US 60/216,547

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/218,232

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/220,112

; PRIOR FILING DATE: 2000-07-21

; PRIOR APPLICATION NUMBER: US 60/221,839

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: PERL Program

; SEQ ID NO 27

; LENGTH: 2180

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20040053258A1 7475603CD1

US-10-332-447-27

Query Match

99.9%; Score 10883; DB 15; Length 2180;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFNKPLPSAGTVPWLQGLICNVNNTCFPOLTPGEEPRGRJNSFNDSLSVRL 60
DB 81 PPLEHHECHFNKPLPSAGTVPWLQGLICNVNNTCFPOLTPGEEPRGRJNSFNDSLSVRL 140
QY 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOQPQTQSPLEPPMLDVAELLTSLRT 120
DB 141 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOQPQTQSPLEPPMLDVAELLTSLRT 200
QY 121 ESLGLALQAOEPPLHSLLEAAEDLAQELLARLSLVELRALLQRPRTGSPLELLSEALCS 180
DB 201 ESLGLALQAOEPPLHSLLEAAEDLAQELLARLSLVELRALLQRPRTGSPLELLSEALCS 260
QY 181 VRGSPSTVGPSNLWYEASDLMELVQBPESALPDSSSLSPACSELIGALDHPHLSLLWRR 240
DB 261 VRGSPSTVGPSNLWYEASDLMELVQBPESALPDSSSLSPACSELIGALDHPHLSLLWRR 320
QY 241 LKPLILGKLLFAPDPTPTTRKLMQAVNRFTFELTLRDVREVMELGPRIFTPMNDSSNVA 300
DB 321 LKPLILGKLLFAPDPTPTTRKLMQAVNRFTFELTLRDVREVMELGPRIFTPMNDSSNVA 380
QY 301 MLQRLQWQDEGRQRPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGLTGRVTEC 360
DB 381 MLQRLQWQDEGRQRPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGLTGRVTEC 440
QY 361 LSLDKLEAAPSEAAALVSRALQLLAEHFRWAGVVFGLPDSDDSPTEHPTDLPDQPGHVRKI 420
DB 441 LSLDKLEAAPSEAAALVSRALQLLAEHFRWAGVVFGLPDSDDSPTEHPTDLPDQPGHVRKI 500
QY 421 RMDIDVTRTNKIRDRFWDPCPAADPLTDLRVWGGFVYLQDLVERAAVRVLSGANPRAG 480
DB 501 RMDIDVTRTNKIRDRFWDPCPAADPLTDLRVWGGFVYLQDLVERAAVRVLSGANPRAG 560
QY 481 LYLAQMPYPYCVDDVFLRVLSRSLPLFLTLLAWIYSVTITVKAVEREKETRLDRTRAMGL 540
DB 561 LYLAQMPYPYCVDDVFLRVLSRSLPLFLTLLAWIYSVTITVKAVEREKETRLDRTRAMGL 620
QY 541 SRAVLNIGWFLSCIGPFLLSAALLVLVLKGDILPYSHPGVVVFLPLAFAVATVTSFLL 600
DB 621 SRAVLNIGWFLSCIGPFLLSAALLVLVLKGDILPYSHPGVVVFLPLAFAVATVTSFLL 680
QY 601 SAFFSRANLAAACGLAYFSLYLPVLCVAVNRDLPAAGRVAASLLSPVAFGFCESLAL 660
DB 681 SAFFSRANLAAACGLAYFSLYLPVLCVAVNRDLPAAGRVAASLLSPVAFGFCESLAL 740
QY 661 LEEQEGQAQMHNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 720
DB 741 LEEQEGQAQMHNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 800
QY 721 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVBEAPPLSPGVSVRSLEKRFPGSPQALRG 780
DB 801 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVBEAPPLSPGVSVRSLEKRFPGSPQALRG 860
QY 781 LSLDPYQGHITAFIYGHNGAGKTTTSLISGLFPSPGGSATFLGHHDVRSMAAIRPHLVC 840
DB 861 LSLDPYQGHITAFIYGHNGAGKTTTSLISGLFPSPGGSATFLGHHDVRSMAAIRPHLVC 920
QY 841 PQYNVLFDMLTVDHEVWVFYGRKGLSAAVVGPEODRLQDVLGVSQSVQTRHLSGGMQR 900
DB 921 PQYNVLFDMLTVDHEVWVFYGRKGLSAAVVGPEODRLQDVLGVSQSVQTRHLSGGMQR 980
QY 901 KLSVAIAIPVGGSQVVIDEPTAGVDPPASRRGIWELLKYREGRTLILSTHLLDEALLGD 960
DB 981 KLSVAIAIPVGGSQVVIDEPTAGVDPPASRRGIWELLKYREGRTLILSTHLLDEALLGD 1040
QY 961 RVAVVAGGRGCCGSPFLFRHLGSGYYLTLVAKARLPLTTNEKADTDNEGSDVTRQEKKN 1020
DB 1041 RVAVVAGGRGCCGSPFLFRHLGSGYYLTLVAKARLPLTTNEKADTDNEGSDVTRQEKKN 1100
QY 1021 GSGQSRVGTGPOLLALVQHVWVFGARLVBELPHELVLVLPYTGAGHDSFATLFRDLRLAE 1080

Db 1101 GSQSRVGTQALLALVOHWVPGARLVEELPHELVLVLPYTGADGSEFATLFRDLDTLAE 1160
Qy 1081 LRLTGYGISTSLBEIFLKVVVECAADTMEDGSCQHLCGTAGLDVTLRLKMPPOETA 1140
Db 1161 LRLTGYGISTSLBEIFLKVVVECAADTMEDGSCQHLCGTAGLDVTLRLKMPPOETA 1220
Qy 1141 LENGEPAGSAPETDQSGSDPAVGRVQGWALTRQQLQALLKRFLLARRSRGLFAQIVLP 1200
Db 1221 LENGEPAGSAPETDQSGSDPAVGRVQGWALTRQQLQALLKRFLLARRSRGLFAQIVLP 1280
Qy 1201 ALFVGLALVLSLIVPPFGHPALRLSPMTYGAQVSFFSEADPGDGRARLLEALLQAGL 1260
Db 1281 ALFVGLALVLSLIVPPFGHPALRLSPMTYGAQVSFFSEADPGDGRARLLEALLQAGL 1340
Qy 1261 BEPPVQHSSHRFSAPEVAEPAEYAKVLAGSNMTPESPSPACQSCQPGARRLLPDCPAAAGGP 1320
Db 1341 BEPPVQHSSHRFSAPEVAEPAEYAKVLAGSNMTPESPSPACQSCQPGARRLLPDCPAAAGGP 1400
Qy 1321 PPOAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGRDPGL 1380
Db 1401 PPOAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGRDPGL 1460
Qy 1381 PSQELGRSVEELWALLSPUPGGALDRVLKNTAWAHSLLDAQDSLKIWFNNKGWHSVAF 1440
Db 1461 PSQELGRSVEELWALLSPUPGGALDRVLKNTAWAHSLLDAQDSLKIWFNNKGWHSVAF 1520
Qy 1441 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLESAALMASSVDVLVSIQVVFAM 1500
Db 1521 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLESAALMASSVDVLVSIQVVFAM 1580
Qy 1501 SFVPASFTLVLIERVTRAKHLQMGSLPFLYHGNFLNMDMCMNLYVPACIVVLIPLAFQ 1560
Db 1581 SFVPASFTLVLIERVTRAKHLQMGSLPFLYHGNFLNMDMCMNLYVPACIVVLIPLAFQ 1640
Qy 1561 QRAVVAANIPALLLLLLLGVMSITPLMYPASFFSPSTAYVVLTCINLFIGINGSWAT 1620
Db 1641 QRAVVAANIPALLLLLLLGVMSITPLMYPASFFSPSTAYVVLTCINLFIGINGSWAT 1700
Qy 1621 FVLELFSQDLQOEVSRILKQVFLIPPHFCFLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1680
Db 1701 FVLELFSQDLQOEVSRILKQVFLIPPHFCFLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1760
Qy 1681 WEVVGKLLAMVIOGPIFLFTLLOHRSQLLPOPRVRSPLLGEDEEDVAREERVVQG 1740
Db 1761 WEVVGKLLAMVIOGPIFLFTLLOHRSQLLPOPRVRSPLLGEDEEDVAREERVVQG 1820
Qy 1741 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCECFGLLVNGAGKTSFTFRWVTGDTLA 1800
Db 1821 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCECFGLLVNGAGKTSFTFRWVTGDTLA 1880
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1860
Db 1881 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1940
Qy 1861 GSGLARLGLSWYADRPAGTYSGGNKRKLATALVGPAPVVFDEPTTGMDDPSARRFLWN 1920
Db 1941 GSGLARLGLSWYADRPAGTYSGGNKRKLATALVGPAPVVFDEPTTGMDDPSARRFLWN 2000
Qy 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 2001 SLLAVVREGRSVMLTSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRV 2060
Qy 1981 PAARSQAAAAFVAAEFPGSELRREAHGGRRLRQLPFGGRCALARVFGELAVHGAEBHGVEDF 2040
Db 2061 PAARSQAAAAFVAAEFPGSELRREAHGGRRLRQLPFGGRCALARVFGELAVHGAEBHGVEDF 2120
Qy 2041 SVSQTMLBEVFLYFSKQGDDETEOKEAGVVDPAQGLQHPKRVSQFOLDPSTAEVTL 2100
Db 2121 SVSQTMLBEVFLYFSKQGDDETEOKEAGVVDPAQGLQHPKRVSQFOLDPSTAEVTL 2180

; Sequence 2, Application US/09858194
; Patent No. US20020061590A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, MARIA
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF
; FILE REFERENCE: MNI-153
; CURRENT APPLICATION NUMBER: US/09/858,194
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/204,211
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-194-2

Query Match 99.9%; Score 10880; DB 9; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPLEHECHFPNKPPLPSAGTVPMLOGLICNVNNTCFPQLTGPEPGRLSNFNDSLSRLL 60
Db 45 PPLEHECHFPNKPPLPSAGTVPMLOGLICNVNNTCFPQLTGPEPGRLSNFNDSLSRLL 104
Qy 61 ADARTVLGGASAHRTLAGLKLATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 120
Db 105 ADARTVLGGASAHRTLAGLKLATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 164
Qy 121 ESGLALQAOQEPHLSLENAEDLAQELLALRSIVELRALLQPRGTSGPLELSEALCS 180
Db 165 ESGLALQAOQEPHLSLENAEDLAQELLALRSIVELRALLQPRGTSGPLELSEALCS 224
Qy 181 VRGFSSTVGPLNMYEASDLMELVQBPESALPDSSLPACSELIGALDHPHLSRLWRR 240
Db 225 VRGFSSTVGPLNMYEASDLMELVQBPESALPDSSLPACSELIGALDHPHLSRLWRR 284
Qy 241 LKPLIILKGLLPADPTPTTRKLMAQVNRTPBELLLRDVREVMELGPRIFTFMDSSNVA 300
Db 285 LKPLIILKGLLPADPTPTTRKLMAQVNRTPBELLLRDVREVMELGPRIFTFMDSSNVA 344
Qy 301 MLQRLIQMDQEGRRQPRGGRDHEALRSFLDPGSGYSQDAHADVGHVGLGRVTEC 360
Db 345 MLQRLIQMDQEGRRQPRGGRDHEALRSFLDPGSGYSQDAHADVGHVGLGRVTEC 404
Qy 361 LSLLDKLEAAPSEALVSRAQLLAHFRFWAGVVFPLGPESSDPTHEPTDPLGPGHVRKI 420
Db 405 LSLLDKLEAAPSEALVSRAQLLAHFRFWAGVVFPLGPESSDPTHEPTDPLGPGHVRKI 464
Qy 421 RMDIDVVRTNKIRDRFPDGPAAADPLTDLRYVVGGVFYLODLVERAAVRLSCANPRAG 480
Db 465 RMDIDVVRTNKIRDRFPDGPAAADPLTDLRYVVGGVFYLODLVERAAVRLSCANPRAG 524
Qy 481 LYLQOMPYPYCVDDVFLVLSRSLPLFLTLAWLYSVTLTKAVVREKETRLRDTMRAMGL 540
Db 525 LYLQOMPYPYCVDDVFLVLSRSLPLFLTLAWLYSVTLTKAVVREKETRLRDTMRAMGL 584
Qy 541 SRVAVLWLGWFLSCIGLPPLLSAAALLVVLKLGDLIPYSHPGVVFLLAAFAVATVTSFLL 600
Db 585 SRVAVLWLGWFLSCIGLPPLLSAAALLVVLKLGDLIPYSHPGVVFLLAAFAVATVTSFLL 644
Qy 601 SAFPFRANLAAACGLAYFSLYLPYVLCVAVNRDLRPPAGGRVAASLLSPVAFGFCESLAL 660
Db 645 SAFPFRANLAAACGLAYFSLYLPYVLCVAVNRDLRPPAGGRVAASLLSPVAFGFCESLAL 704
Qy 661 LEQGEQAQHNHVTREPTADVFSIAQVSGLLDAAALYGLATWYLEAVCPGQYGIPEPMN 720
Db 705 LEQGEQAQHNHVTREPTADVFSIAQVSGLLDAAALYGLATWYLEAVCPGQYGIPEPMN 764
Qy 721 FPFRRSYWCGRPPKSPAPCFTPLDPKVLVEEAPPGLSPGVSVRSLEKFPFGSPQALRG 780

Db 765 FFRERSYMCPRPKSPAPCTPLDPKVLVEAPPGLSPGVSVRSLEKRPFGSPQPALRG 824
Qy 781 LSLDFVQGHITAFIHLNGAGKTTTILSISLGLPPSGGSFALIGHDVRSSMAIRPHLVC 840
Db 825 LSLDFYQGHITAFIHLNGAGKTTTILSISLGLPPSGGSFALIGHDVRSSMAIRPHLVC 884
Qy 841 POYNVLFDMITVDEHWFYGRILKGLSAAVVGPEQDRLLQDVGLSKVQSVQTRHLSGGMR 900
Db 885 POYNVLFDMITVDEHWFYGRILKGLSAAVVGPEQDRLLQDVGLSKVQSVQTRHLSGGMR 944
Qy 901 KLSVAIAFVGGQVILDEPTAGVDPASRRGIWELLKYREGRTILISTHLLDEALLGD 960
Db 945 KLSVAIAFVGGQVILDEPTAGVDPASRRGIWELLKYREGRTILISTHLLDEALLGD 1004
Qy 961 RVAVVAGRLCCCGSPLFRLRHLSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN 1020
Db 1005 RVAVVAGRLCCCGSPLFRLRHLSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN 1064
Qy 1021 QSGSRVGTPLQALVQHWVPGARLVEELPHELVLVLVLYPYTGADHGSFATLFRLELTRLAE 1080
Db 1065 QSGSRVGTPLQALVQHWVPGARLVEELPHELVLVLVLYPYTGADHGSFATLFRLELTRLAE 1124
Qy 1081 LRLTGYGSDTSLBEIFLKVVECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1140
Db 1125 LRLTGYGSDTSLBEIFLKVVECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1184
Qy 1141 LENGEPAGSAPETDQSGPDVGRVQGWALTRQQLQALLKRFLLARESRGLFAQIVLP 1200
Db 1185 LENGEPAGSAPETDQSGPDVGRVQGWALTRQQLQALLKRFLLARESRGLFAQIVLP 1244
Qy 1201 ALFVGLALVFSILVPPFGHYPALRLSPMTWYGAQVFFSEDPGDPGRARLLLEALQEAGL 1260
Db 1245 ALFVGLALVFSILVPPFGHYPALRLSPMTWYGAQVFFSEDPGDPGRARLLLEALQEAGL 1304
Qy 1261 BEPPVQSHSRPSADEVPAEVAKVILASGNWTPESPSPACQSQOPGARILLPDCPAAAGP 1320
Db 1305 BEPPVQSHSRPSADEVPAEVAKVILASGNWTPESPSPACQSQOPGARILLPDCPAAAGP 1364
Qy 1321 PPPQAVTSGEVQNLTRGNLSDFLVKTYPRLVROGLTKKWNVEVRYGGSFSGRDRPGL 1380
Db 1365 PPPQAVTSGEVQNLTRGNLSDFLVKTYPRLVROGLTKKWNVEVRYGGSFSGRDRPGL 1424
Qy 1381 PSGQELGRSVEELWALLSPLFGGALDRVLKNTLAWAHSIDQAQDSLKIWFNKGHSMVAF 1440
Db 1425 PSGQELGRSVEELWALLSPLFGGALDRVLKNTLAWAHSIDQAQDSLKIWFNKGHSMVAF 1484
Qy 1441 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVVPAM 1500
Db 1485 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVVPAM 1544
Qy 1501 SFVPASFTVLVLEERVTRAKHLQLMGGISPTLYLGNFLWDMCNVLPACIVVLIFLAFQ 1560
Db 1545 SFVPASFTVLVLEERVTRAKHLQLMGGISPTLYLGNFLWDMCNVLPACIVVLIFLAFQ 1604
Qy 1561 QRAYVAPANLPALELLLYGWSITPLMYPASFVSPSTAYVVLTCINLFIGINGSNAT 1620
Db 1605 QRAYVAPANLPALELLLYGWSITPLMYPASFVSPSTAYVVLTCINLFIGINGSNAT 1664
Qy 1621 FVLELFSQKLOEVSRIILKQVFLIPPHFCLOGLIDMVRNQAMADAFRLGDROFQSPLR 1680
Db 1665 FVLELFSQKLOEVSRIILKQVFLIPPHFCLOGLIDMVRNQAMADAFRLGDROFQSPLR 1724
Qy 1681 WEVVGKLLAMVIOQPLFLFTLLLOHRSQLLPOPRVRSPLPLGDEEDVARERERVQOG 1740
Db 1725 WEVVGKLLAMVIOQPLFLFTLLLOHRSQLLPOPRVRSPLPLGDEEDVARERERVQOG 1784
Qy 1741 ATQGDVLVRLNLTKVYRGORMPAVDRCLGIPPGCEFCGLLVNGAGKTSITFRMTVGTDLA 1800
Db 1785 ATQGDVLVRLNLTKVYRGORMPAVDRCLGIPPGCEFCGLLVNGAGKTSITFRMTVGTDLA 1844
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLURGVPEAQVQTA 1860
Db 1845 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLURGVPEAQVQTA 1904

Qy 1861 GSGLARILGSLWYADRPACTYSGNKRKLATALLVGDPAVVPFLDEPTTCMDPSARRFLWN 1920
Db 1905 GSGLARILGSLWYADRPACTYSGNKRKLATALLVGDPAVVPFLDEPTTCMDPSARRFLWN 1964
Qy 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIWNGRFRCLASPOHLKGRFAAGHTLTLRV 1980
Db 1965 SLLAVVREGRSVMLTSHSMECEALCSRLAIWNGRFRCLASPOHLKGRFAAGHTLTLRV 2024
Qy 1981 PAARSQAAAAFVAABFFPGSELREAHGGRLRFQLPPGRCALARVFGELAVHGAHGVDF 2040
Db 2025 PAARSQAAAAFVAABFFPGSELREAHGGRLRFQLPPGRCALARVFGELAVHGAHGVDF 2084
Qy 2041 SVSQTMLTEEVFLYFSKQGDSDTEOEKAGVGVDPAPGLQHPKXVSQFLDDPSTAEVTL 2100
Db 2085 SVSQTMLTEEVFLYFSKQGDSDTEOEKAGVGVDPAPGLQHPKXVSQFLDDPSTAEVTL 2144
RESULT 6
US-10-154-419-2
; Sequence 2, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8039, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-2
Query Match 99.9%; Score 10880; DB 14; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPLEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCFPOLTPGEBFGRLSNFNDSLVSRLL 60
Db 45 PPLEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCFPOLTPGEBFGRLSNFNDSLVSRLL 104
Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLIRT 120
Db 105 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLIRT 164
Qy 121 ESGLGLAQOEPLHSLLEAAEDLAQELLARLSVELRALLQRPRTSGPSELLSEALCS 180
Db 165 ESGLGLAQOEPLHSLLEAAEDLAQELLARLSVELRALLQRPRTSGPSELLSEALCS 224
Qy 181 VRGSPSTVGSPLNMYEASDLMLVQBPESALPOSSLS PACSELIGALDHPRLRWR 240
Db 225 VRGSPSTVGSPLNMYEASDLMLVQBPESALPOSSLS PACSELIGALDHPRLRWR 284
Qy 241 LKPLILGKLLFAPDTPPTRKLMAQVNTFEELTLRDVREVMELGPRIFTFMDNSSVA 300
Db 285 LKPLILGKLLFAPDTPPTRKLMAQVNTFEELTLRDVREVMELGPRIFTFMDNSSVA 344
Qy 301 MLQRLQWQDGRQRPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVTGLGRVTEC 360
Db 345 MLQRLQWQDGRQRPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVTGLGRVTEC 404
Qy 361 LSLDKLEAAPSEAAALVSRALQLLAEHRFWAGVVFPGEDSSDPTHEHPTDLPDGPGHVRIKI 420
Db 405 LSLDKLEAAPSEAAALVSRALQLLAEHRFWAGVVFPGEDSSDPTHEHPTDLPDGPGHVRIKI 464

QY 421 RMDIDVTRTNKIRDRFWDPCGPAADPLTDRIYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
DB 465 RMDIDVTRTNKIRDRFWDPCGPAADPLTDRIYVWGGFVYLQDLVERAAVRVLSGANPRAG 524
QY 481 LYLOQMPYPCVDDVFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDRTMRAMGL 540
DB 525 LYLOQMPYPCVDDVFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDRTMRAMGL 584
QY 541 SRAVLWLGWFLSLCLGPFLLSAAILLVVLKLGDIILPYSHPGVVFLLFAFAVAIVTQSFLL 600
DB 585 SRAVLWLGWFLSLCLGPFLLSAAILLVVLKLGDIILPYSHPGVVFLLFAFAVAIVTQSFLL 644
QY 601 SAFPSRANLAAACGLAYFSLYLPIVLCVWRDLRPPAGGRVAASLLSPVAFGFCESLAL 660
DB 645 SAFPSRANLAAACGLAYFSLYLPIVLCVWRDLRPPAGGRVAASLLSPVAFGFCESLAL 704
QY 661 LEEQGEAQNHNCTRTADVSLAOSVGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 720
DB 705 LEEQGEAQNHNCTRTADVSLAOSVGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 764
QY 721 FPFRRSYWCGRPPKSPAPCTPLDPKVLVEEAPPGLSPGVSRLEKRPFGSPQPALRG 780
DB 765 FPFRRSYWCGRPPKSPAPCTPLDPKVLVEEAPPGLSPGVSRLEKRPFGSPQPALRG 824
QY 781 LSLDFYQCHITAFILGHNGAGKTTLSILSGLFPSPGSGSAFILGHDRSSMAAIRPHLGV 840
DB 825 LSLDFYQCHITAFILGHNGAGKTTLSILSGLFPSPGSGSAFILGHDRSSMAAIRPHLGV 884
QY 841 PQYNVLFMDLTVDBHWFYGRKGLSAAVGPEDRLLDVGLVSKOSVOTRHLSSGMQR 900
DB 885 PQYNVLFMDLTVDBHWFYGRKGLSAAVGPEDRLLDVGLVSKOSVOTRHLSSGMQR 944
QY 901 KLSVAIAFVGSQVVIIDEPAGVDPASRRGIWELLKRYREGRTLILSTHLLDEAELLGD 960
DB 945 KLSVAIAFVGSQVVIIDEPAGVDPASRRGIWELLKRYREGRTLILSTHLLDEAELLGD 1004
QY 961 RVAVVAGRLCCCSPLFLRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTROBKN 1020
DB 1005 RVAVVAGRLCCCSPLFLRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTROBKN 1064
QY 1021 GSQSRVGTTPOLLALVOHWPVGAIRLVEELPHELVLVLPYTGADHGSFATLFLRELDTRLAE 1080
DB 1065 GSQSRVGTTPOLLALVOHWPVGAIRLVEELPHELVLVLPYTGADHGSFATLFLRELDTRLAE 1124
QY 1081 LRLTGYGSDTSLEEIFLKVVEECAADTDMEDGSCGQHLCTGTGIAGLDVTLLRKMPPQETA 1140
DB 1125 LRLTGYGSDTSLEEIFLKVVEECAADTDMEDGSCGQHLCTGTGIAGLDVTLLRKMPPQETA 1184
QY 1141 LENGEPAAGSAPETDQSGPDVAVRQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1200
DB 1185 LENGEPAAGSAPETDQSGPDVAVRQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1244
QY 1201 ALFVGLALVFLSLIYVPPGHYPALRLSPMTYGAQVSFFSEADAPGPGARLLEALLQAGL 1260
DB 1245 ALFVGLALVFLSLIYVPPGHYPALRLSPMTYGAQVSFFSEADAPGPGARLLEALLQAGL 1304
QY 1261 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNMTPEPSPACQCSQPGARLLPDCPAAAGGP 1320
DB 1305 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNMTPEPSPACQCSQPGARLLPDCPAAAGGP 1364
QY 1321 PPOQAVTGSBQVWQNLTRNLSDFLVKTYPRLVQGLKTKKWNVEVYGGFSLGGRDPGL 1380
DB 1365 PPOQAVTGSBQVWQNLTRNLSDFLVKTYPRLVQGLKTKKWNVEVYGGFSLGGRDPGL 1424
QY 1381 PSQGLGRSVEELWALLSPLPGGALDRVLKNTAWAHSLLDAQDSLKIWFNNKWHSMVAF 1440
DB 1425 PSQGLGRSVEELWALLSPLPGGALDRVLKNTAWAHSLLDAQDSLKIWFNNKWHSMVAF 1484
QY 1441 VNRSASNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSTCVVPAM 1500
DB 1485 VNRSASNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSTCVVPAM 1544

QY 1501 SFVPASFTVLIERVTRAKHLOLMGCLSPTLVWLNFLWDMCNLYLPACIVLILFAFQ 1560
DB 1545 SFVPASFTVLIERVTRAKHLOLMGCLSPTLVWLNFLWDMCNLYLPACIVLILFAFQ 1604
QY 1561 QRAYVAPANUPALLLLLLYGSITPLMYPASFPFSPSTAYVVLTCINILFIGINGSMAT 1620
DB 1605 QRAYVAPANUPALLLLLLYGSITPLMYPASFPFSPSTAYVVLTCINILFIGINGSMAT 1664
QY 1621 FVLELFSQDKLQVSRILKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGDROFQSPLR 1680
DB 1665 FVLELFSQDKLQVSRILKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGDROFQSPLR 1724
QY 1681 WEVVGKLLAMVIOGFLFLLTLLQHRSQLLPQPRVRSILPILGEEDVARERERVQOG 1740
DB 1725 WEVVGKLLAMVIOGFLFLLTLLQHRSQLLPQPRVRSILPILGEEDVARERERVQOG 1784
QY 1741 ATQGDVLVLRNLTKVYRGQMPAVDRLCLGIPGECFGLLVNGAGKTSIFRMVGTDLA 1800
DB 1785 ATQGDVLVLRNLTKVYRGQMPAVDRLCLGIPGECFGLLVNGAGKTSIFRMVGTDLA 1844
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1860
DB 1845 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1904
QY 1861 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDSARRFLWN 1920
DB 1905 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDSARRFLWN 1964
QY 1921 SLAVVREGSRVMTLSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLV 1980
DB 1965 SLAVVREGSRVMTLSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLV 2024
QY 1981 PAARSQPAFAVAAEPFSGSELREAHGRLRFQLPFGRCALARVFGELAVHGAHGVEDF 2040
DB 2025 PAARSQPAFAVAAEPFSGSELREAHGRLRFQLPFGRCALARVFGELAVHGAHGVEDF 2084
QY 2041 SVSQTMLVEVLYFSKQDDETEOKEAGVVDPAQGLQHPKRVSOQLDDPSTAEVTVL 2100
DB 2085 SVSQTMLVEVLYFSKQDDETEOKEAGVVDPAQGLQHPKRVSOQLDDPSTAEVTVL 2144

RESULT 7

US-10-618-281-44
; Sequence 44, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-44

Query Match 99.8%; Score 10874.5; DB 17; Length 2147;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 PPLEHHECHFPNKPPLPSAGTVPMVLOGLICNVNNTCFPQLTPGEBPGRLSFNDSLSRLL 60
DB 47 PPLEHHECHFPNKPPLPSAGTVPMVLOGLICNVNNTCFPQLTPGEBPGRLSFNDSLSRLL 106

QY 61 ADARTVLGASAHRTLAGLKIATLRAARSTAQPOPTKQSPLEPPMLDVAELTSLRT 120
DB 107 ADARTVLGASAHRTLAGLKIATLRAARSTAQPOPTKQSPLEPPMLDVAELTSLRT 166
QY 121 ESLGALGOAEPHLSLEAAEDIAQELLALRSVELRALLORPCTSGPLELSEALCS 180
DB 167 ESLGALGOAEPHLSLEAAEDIAQELLALRSVELRALLORPCTSGPLELSEALCS 226
QY 181 VRGSPSTVGPNSLWYEAADLMELVGOEPESALPDSLSLSPACSELIGALDHSPLSRLWRR 240
DB 227 VRGSPSTVGPNSLWYEAADLMELVGOEPESALPDSLSLSPACSELIGALDHSPLSRLWRR 286
QY 241 LKPLILGKLLFAPDTPFTFKLMAQVNRTEBELTLRDVREWEMLGPRIFTFMDSSNVA 300
DB 287 LKPLILGKLLFAPDTPFTFKLMAQVNRTEBELTLRDVREWEMLGPRIFTFMDSSNVA 346
QY 301 MLQRLQWODEGRRQPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGTIGRVTEC 360
DB 347 MLQRLQWODEGRRQPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGTIGRVTEC 406
QY 361 LSLDKLEAAPGEAALVSRALQLLAHRFPWAGVVFGLGEDSDPTBHTPDLPBGHVRKI 420
DB 407 LSLDKLEAAPGEAALVSRALQLLAHRFPWAGVVFGLGEDSDPTBHTPDLPBGHVRKI 466
QY 421 RMDIDVTRTKIRDRFWDPCPADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTKIRDRFWDPCPADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWISVTLTVKAVVREKTRLDTRWAMGL 540
DB 527 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWISVTLTVKAVVREKTRLDTRWAMGL 586
QY 541 SRAVLWLGWFLSCLGPFLLSAAILVLKLGIDILPYSHPGVVFLFLAFAFVATVTSPLL 600
DB 587 SRAVLWLGWFLSCLGPFLLSAAILVLKLGIDILPYSHPGVVFLFLAFAFVATVTSPLL 646
QY 601 SAFPSTRANLAAACGLAYFSLYLPVYLVCVWRDLRDPAGGRVAASLLSPVAFGFCESIAL 660
DB 647 SAFPSTRANLAAACGLAYFSLYLPVYLVCVWRDLRDPAGGRVAASLLSPVAFGFCESIAL 706
QY 661 LEEQEGQWQHNVTGTRPTADVFLSAQVSGLLILDAALYGLATWILEAVCPQOYGIPEPWN 720
DB 707 LEEQEGQWQHNVTGTRPTADVFLSAQVSGLLILDAALYGLATWILEAVCPQOYGIPEPWN 766
QY 721 PPFRRSYWCGPRPKSPACPTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPQPALRG 780
DB 767 PPFRRSYWCGPRPKSPACPTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPQPALRG 826
QY 781 LSLDFYQGHITAFIHNAGKTTTILSILSGLFPSPGSAFILGHDRSSMAAIRPHLGVC 840
DB 827 LSLDFYQGHITAFIHNAGKTTTILSILSGLFPSPGSAFILGHDRSSMAAIRPHLGVC 886
QY 841 POYNVLPMLTVDHEWVYGBRLKGLSAVAVGPEQDRLLQDVLVSKSVQVTRHLSGGMQR 900
DB 887 POYNVLPMLTVDHEWVYGBRLKGLSAVAVGPEQDRLLQDVLVSKSVQVTRHLSGGMQR 946
QY 901 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIIWELLKLYREGRTLILSTHLDABEALLGD 960
DB 947 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIIWELLKLYREGRTLILSTHLDABEALLGD 1006
QY 961 RVAVVAGRLCCCGSPFLRRLHLSGGYLLTLVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
DB 1007 RVAVVAGRLCCCGSPFLRRLHLSGGYLLTLVKARLPLTTNEKADTMEGSDVTRQEKKN 1066
QY 1021 GSQSRVGTGPOLLALVQHWVPGARLVEBELPHELVLVLPYTGANDGSPATLRFELDTRLAE 1080
DB 1067 GSQSRVGTGPOLLALVQHWVPGARLVEBELPHELVLVLPYTGANDGSPATLRFELDTRLAE 1126
QY 1081 LRLTYGIDSFTSLEBIFLKVVEECAAADTMDGSGCGHLCGTGAGLDVTLRLKMPPOBTA 1140
DB 1127 LRLTYGIDSFTSLEBIFLKVVEECAAADTMDGSGCGHLCGTGAGLDVTLRLKMPPOBTA 1186
QY 1141 LENGEPAGSAPETDQSGSPDAVGRVQGWALTRQOLQALLKRLFLARRRRRGLFAQIVLP 1200

DB 1187 LENGEPAGSAPETDQSGSPDAVGRVQGWALTRQOLQALLKRLFLARRRRRGLFAQIVLP 1246
QY 1201 ALFVGLALVFSLIYVPPFGHYPALRLSPMYGNAQVVSFFSEDAPGDPGRARLLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSLIYVPPFGHYPALRLSPMYGNAQVVSFFSEDAPGDPGRARLLLEALLQEAGL 1306
QY 1261 BEPPVQSHSHRFSAPAEVPAEVAKVLASGNWTPESPSACQSQOPCARLLPDCPAAAGP 1320
DB 1307 BEPPVQSHSHRFSAPAEVPAEVAKVLASGNWTPESPSACQSQOPCARLLPDCPAAAGP 1366
QY 1321 PPPQAVTSGSEVQNLGTRNLSDFLVKTYPRLVQGLTKKQVNEVRYGGFSLGGRDPGL 1380
DB 1367 PPPQAVTSGSEVQNLGTRNLSDFLVKTYPRLVQGLTKKQVNEVRYGGFSLGGRDPGL 1426
QY 1381 PSGQELGRSVEBELWALLSPLPGGALDRVLKNTAWAHSLDQAQDSLKIWFNNKGHSMVAF 1440
DB 1427 PSGQELGRSVEBELWALLSPLPGGALDRVLKNTAWAHSLDQAQDSLKIWFNNKGHSMVAF 1486
QY 1441 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVTFAM 1500
DB 1487 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVTFAM 1546
QY 1501 SFVPASFTVLVIERBVTTRAKHLQMGGLSPTLYWLGNFPLWDMCNLYVPACIVVLIFLAFQ 1560
DB 1547 SFVPASFTVLVIERBVTTRAKHLQMGGLSPTLYWLGNFPLWDMCNLYVPACIVVLIFLAFQ 1606
QY 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSPSTAYVVLTCNLNFIGINGSWAT 1620
DB 1607 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSPSTAYVVLTCNLNFIGINGSWAT 1666
QY 1621 FVLELFSQDQKQLEVSRLKQVFLIFPHFCILGRGLIDMVRNOAMADAPERLGDROFQSP 1679
DB 1667 FVLELFSQDQKQLEVSRLKQVFLIFPHFCILGRGLIDMVRNOAMADAPERLGDROFQSP 1726
QY 1680 RWEVVGKXLLAMVIOGPFLLFTLLQHRSQLLPQPRVRSPLLGEEDVVARERERVVQ 1739
DB 1727 RWEVVGKXLLAMVIOGPFLLFTLLQHRSQLLPQPRVRSPLLGEEDVVARERERVVQ 1786
QY 1740 GATQGDVLVLRNLTKVYRGQRMVADRLCLGIPGCEGGLGVNGAGKSTFRMVVTGDTL 1799
DB 1787 GATQGDVLVLRNLTKVYRGQRMVADRLCLGIPGCEGGLGVNGAGKSTFRMVVTGDTL 1846
QY 1800 ASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVAQT 1859
DB 1847 ASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVAQT 1906
QY 1860 AGSGLARLGLSWYADRPAQTYSGGNKRLATATALVGDPAVVFLEDEPTTGMDPSARRFLW 1919
DB 1907 AGSGLARLGLSWYADRPAQTYSGGNKRLATATALVGDPAVVFLEDEPTTGMDPSARRFLW 1966
QY 1920 NSLLAVVREGSVMLTSHSMECEALCSRLAIMYNGRPRCLGSPQHLKGRFAAGHTLTLR 1979
DB 1967 NSLLAVVREGSVMLTSHSMECEALCSRLAIMYNGRPRCLGSPQHLKGRFAAGHTLTLR 2026
QY 1980 VPAARSPAAAFAAEFPFSGELREAHGRLRFQPPGRCALARVFGELAVHGAHGVED 2039
DB 2027 VPAARSPAAAFAAEFPFSGELREAHGRLRFQPPGRCALARVFGELAVHGAHGVED 2086
QY 2040 FSVSQTMLVEEVLYFSKQDQKDETEBEQKEAGVGDVPAPGLQHPKRVVSQFLLDDPSTASTV 2099
DB 2087 FSVSQTMLVEEVLYFSKQDQKDETEBEQKEAGVGDVPAPGLQHPKRVVSQFLLDDPSTASTV 2146
QY 2100 L 2100
DB 2147 L 2147

RESULT 8

US-10-114-270-176

; Sequence 176, Application US/10114270

; Publication No. US20040030110A1

; GENERAL INFORMATION:

APPLICANT: Guo, Xiaojia
APPLICANT: Kekuda, Ramesh
APPLICANT: Miller, Charles E.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Patturajan, Meera
APPLICANT: Liu, Ziaohong
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Li, Li
APPLICANT: Vernet, Corine
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glenna
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shimkets, Richard A.
APPLICANT: Gangolli, Esha A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Caeman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liete, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
TITLE OF INVENTION: No. US20040030110alel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-322C
CURRENT APPLICATION NUMBER: US/10/114,270
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 470
SEQ ID NO 176
LENGTH: 2059
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-176

Query Match 95.4%; Score 10392.5; DB 15; Length 2059;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 PPLEHHECHPNKPLPSAGTVPVWQGLICNVNNTCFPQLTPGEPGRLSNFNSLSVRL 60
Db 47 PPLEHHECHPNKPLPSAGTVPVWQGLICNVNNTCFPQLTPGEPGRLSNFNSLSVRL 106
Qy 61 ADARTVLGGASAHRTLAGLKLATLRAARSTAOPTKOSPLPPMLDVAELLTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLKLATLRAARSTAOPTKOSPLPPMLDVAELLTSLRT 166

Qy 121 ESLGLALGQAQEPHLSLLEAAEDLAQELALRSLVELRALLQRPRTSGPFLLESLALCS 180
Db 167 ESLGLALGQAQEPHLSLLEAAEDLAQELALRSLVELRALLQRPRTSGPFLLESLALCS 226
Qy 181 VRGPSTTVGPSLNWYASDLMELVQEPESALPDSISLSPACSELIGALDHPHLSRLWRR 240
Db 227 VRGPSTTVGPSLNWYASDLMELVQEPESALPDSISLSPACSELIGALDHPHLSRLWRR 286
Qy 241 LKPLILGKLLFAPDTPFTTKLMAQVNRFTFELTLDRDREVWEMLGPRIITFFNDSSNVA 300
Db 287 LKPLILGKLLFAPDTPFTTKLMAQVNRFTFELTLDRDREVWEMLGPRIITFFNDSSNVA 346
Qy 301 MLQRLQMDQEGRRQRRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLGTGRVTEC 360
Db 347 MLQRLQMDQEGRRQRRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLGTGRVTEC 406
Qy 361 LSLDKLEAPSEALVSRALQLLAEHRFWAGVVFGLGPESSDTEHPTDPLGPHGVRIKI 420
Db 407 LSLDKLEAPSEALVSRALQLLAEHRFWAGVVFGLGPESSDTEHPTDPLGPHGVRIKI 466
Qy 421 RMDIDVVTINKIRDRFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAG 480
Db 467 RMDIDVVTINKIRDRFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAG 526
Qy 481 LYLQMPYPCYVDDVFLRVLRSLSPLPLTLAWIYSVTLTAKAVVRKETHRLDTRMAMGL 540
Db 527 LYLQMPYPCYVDDVFLRVLRSLSPLPLTLAWIYSVTLTAKAVVRKETHRLDTRMAMGL 586
Qy 541 SRAVLMLGFLSCIGPFLSAAALLVLKLGDTLPYSHPGVWFLFLAAFAVATVTSQFL 600
Db 587 SRAVLMLGFLSCIGPFLSAAALLVLKLGDTLPYSHPGVWFLFLAAFAVATVTSQFL 646
Qy 601 SAFFSRANLAAACGLAYFSLYPYVLCVNRDLRPAAGRVAASLLSPVAFGCESLAL 660
Db 647 SAFFSRANLAAACGLAYFSLYPYVLCVNRDLRPAAGRVAASLLSPVAFGCESLAL 706
Qy 661 LEEQGEQAQHNVTGTRPTADVLSLAQVSGLLLDALYGLATWYLEAVCGQVGIPEPWN 720
Db 707 LEEQGEQAQHNVTGTRPTADVLSLAQVSGLLLDALYGLATWYLEAVCGQVGIPEPWN 766
Qy 721 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPSPQPALRG 780
Db 767 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPSPQPALRG 826
Qy 781 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGSAPILGHDVRSMAAIRPHLGV 840
Db 827 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGSAPILGHDVRSMAAIRPHLGV 886
Qy 841 PQYNVLFMDLTVDEHWFYGRKLGLSAAVVGPDRLQDVLGVSKQSVQTRHLSGQMOR 900
Db 887 PQYNVLFMDLTVDEHWFYGRKLGLSAAVVGPDRLQDVLGVSKQSVQTRHLSGQMOR 946
Qy 901 KLSVAITAFVGSQVWILDEPTAGVDPASRRGIWELLKRYREGRTLILSTHHLDEAEELGD 960
Db 947 KLSVAITAFVGSQVWILDEPTAGVDPASRRGIWELLKRYREGRTLILSTHHLDEAEELGD 1006
Qy 961 RVAVAGGRLLCCCGSPFLRRHLGSGYLLTVKARPLTTNEKADTDMGSDVTROSKN 1020
Db 1007 RVAVAGGRLLCCCGSPFLRRHLGSGYLLTVKARPLTTNEKADTDMGSDVTROSKN 1066
Qy 1021 GSGQSRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGANDGSPATIFRELDTRLAE 1080
Db 1067 GSGQSRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGANDGSPATIFRELDTRLAE 1126
Qy 1081 LRLTGIGISDTSLEEIFLKVVEECAADTMDGSCGHLCTGIAGLDVTLRLKMPPOETA 1140
Db 1127 LRLTGIGISDTSLEEIFLKVVEECAADTMDGSCGHLCTGIAGLDVTLRLKMPPOETA 1186
Qy 1141 LENGEPAGSAPETDQSGGPDVAGRVQGWALTRQOLQALLKRLFLARRSRGLFAQIVLP 1200
Db 1187 LENGEPAGSAPETDQSGGPDVAGRVQGWALTRQOLQALLKRLFLARRSRGLFAQIVLP 1246

QY 1201 ALFVGLALVFSIIVPPFGHYPALRLSPMTMYGAQVSFFSDEADPGDFGRARLLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSIIVPPFGHYPALRLSPMTMYGAQVSFFSDEADPGDFGRARLLLEALLQEAGL 1306
QY 1261 EBPVQVSHRSAPAEVPAEVAKVILASGNWTPESPACQSQCPGARLLLPDCAAGGP 1320
DB 1307 EBPVQVSHRSAPAEVPAEVAKVILASGNWTPESPACQSQCPGARLLLPDCAAGGP 1366
QY 1321 PPQAVTSGEVQVNLTRGNLSDFLVKTYPRLVROGLTKKWNVRYGGSLGRDRPGL 1380
DB 1367 PPQAVTSGEVQVNLTRGNLSDFLVKTYPRLVROGLTKKWNVRYGGSLGRDRPGL 1426
QY 1381 PSQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIWFNKGHSMVAF 1440
DB 1427 PSQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIWFNKGHSMVAF 1486
QY 1441 VNRSANILRAHLPGPAPARHAHSITLHNPLNLTKEQLSEALMASSVDVLVSIICVVFAM 1500
DB 1487 VNRSANILRAHLPGPAPARHAHSITLHNPLNLTKEQLSEALMASSVDVLVSIICVVFAM 1546
QY 1501 SFVPASFTLVLTERRVTRAKHLQMLGGLSPTLYLWGNFLWDMCNVLPACIVLILFLAPQ 1560
DB 1547 SFVPASFTLVLTERRVTRAKHLQMLGGLSPTLYLWGNFLWDMCNVLPACIVLILFLAPQ 1588
QY 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFSPFSTAYVVLTCINLFIGINGSMAT 1620
DB 1589 ----- 1588
QY 1621 FVLELFSQKLOEVSRILKQVFLIPPHFCCLGRGLIDMVRNOMADAFERLGDROQPSPLR 1680
DB 1589 ----- KLQEVSRILKQVFLIPPHFCCLGRGLIDMVRNOMADAFERLGDROQPSPLR 1639
QY 1681 WEVGNKLLAMVIOQPLFLFTLLQHRSQLLPQPRVSLPLGGEEDVARERERVQOG 1740
DB 1640 WEVGNKLLAMVIOQPLFLFTLLQHRSQLLPQPRVSLPLGGEEDVARERERVQOG 1699
QY 1741 ATQGDVLVRLNLTKVYRQMPAVDRCLGIPPGCEGGLGVNGAGKSTTFRMVGTDLA 1800
DB 1700 ATQGDVLVRLNLTKVYRQMPAVDRCLGIPPGCEGGLGVNGAGKSTTFRMVGTDLA 1759
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARLURGVPEAQVQTA 1860
DB 1760 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARLURGVPEAQVQTA 1819
QY 1861 GSGLARLGLSWYADRPAGTYSGGNKRLATALINGDPVAVFLDEPTTGMPSARRFLWN 1920
DB 1820 GSGLARLGLSWYADRPAGTYSGGNKRLATALINGDPVAVFLDEPTTGMPSARRFLWN 1879
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
DB 1880 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1939
QY 1981 PAARSQAAAAFVAABFPQSELREAHGRLRQLPPGRCALARVFGELAVHGAEGHVEDF 2040
DB 1940 PAARSQAAAAFVAABFPQSELREAHGRLRQLPPGRCALARVFGELAVHGAEGHVEDF 1999
QY 2041 SVSQTMLEEVFLYFSKQKDEDEEKEAGVGDVPAPGLOHPKRVSOFLDDPSTAEVTL 2100
DB 2000 SVSQTMLEEVFLYFSKQKDEDEEKEAGVGDVPAPGLOHPKRVSOFLDDPSTAEVTL 2059

RESULT 9
US-10-182-006-4
; Sequence 4, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: PCT/US01/02191

; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1873
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-006-4

Query Match 84.6%; Score 9213.5; DB 15; Length 1873;
Best Local Similarity 97.4%; Pred No. 0;
Matches 1789; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLVSRLL 60
DB 47 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLVSRLL 106
QY 61 ADAPTILGCGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 120
DB 107 ADAPTILGCGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 166
QY 121 ESLGLALGQAQEPPLHSLLEAAEDLAQELLALRSIVELRALIQRPGTSGPLELSEALCS 180
DB 167 ESLGLALGQAQEPPLHSLLEAAEDLAQELLALRSIVELRALIQRPGTSGPLELSEALCS 226
QY 181 VRGFSSTVGPSSLNWYEASDLMEVQGPESALPDSSISPACSELIGALDHSPLSRLWRR 240
DB 227 VRGFSSTVGPSSLNWYEASDLMEVQGPESALPDSSISPACSELIGALDHSPLSRLWRR 286
QY 241 LKPLILGKLLPAPDTPTRKLMQAVNRTFEBLTLLRDVREVWEMLPRIPTFMNDSSNVA 300
DB 287 LKPLILGKLLPAPDTPTRKLMQAVNRTFEBLTLLRDVREVWEMLPRIPTFMNDSSNVA 346
QY 301 MLQRLQWQDGRGRRQPPGGRDHNEALRSFLDPCSGGYSWODAHADVGLVGTIGRVTEC 360
DB 347 MLQRLQWQDGRGRRQPPGGRDHNEALRSFLDPCSGGYSWODAHADVGLVGTIGRVTEC 406
QY 361 LSLDKLEAAPSEAAALVSRAQLQALLAEHRFWAGVFLGPEDDSDPTHEPTDLPQGHVRIKI 420
DB 407 LSLDKLEAAPSEAAALVSRAQLQALLAEHRFWAGVFLGPEDDSDPTHEPTDLPQGHVRIKI 466
QY 421 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYWGGGFYLLQDLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYWGGGFYLLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCVDDVFLRVLSRSLPLFTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL 540
DB 527 LYLOQMPYPCVDDVFLRVLSRSLPLFTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL 586
QY 541 SRAVLWLGWFLSCIGLPPFLLSAALLVLVKGDIILFYSHPGVVFLFLAFAVATVTSFLL 600
DB 587 SRAVLWLGWFLSCIGLPPFLLSAALLVLVKGDIILFYSHPGVVFLFLAFAVATVTSFLL 646
QY 601 SAFFSRANLAAACGGLAYFSLYLPYLVCAWRDLDPAGGRVAAASLLSPVAFGFCESIAL 660
DB 647 SAFFSRANLAAACGGLAYFSLYLPYLVCAWRDLDPAGGRVAAASLLSPVAFGFCESIAL 706
QY 661 LEEQEGQAQHNVTGTRTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPWN 720
DB 707 LEEQEGQAQHNVTGTRTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPWN 766
QY 721 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSVRSVRSLEKRPFGSPQALRG 780
DB 767 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSVRSVRSLEKRPFGSPQALRG 826
QY 781 LSLDFYQGHITAFILGHNGAGKTTTLLSILGILFPSPGSGSAFILLGHVDVRSSMAAIRPHLQVC 840
DB 827 LSLDFYQGHITAFILGHNGAGKTTTLLSILGILFPSPGSGSAFILLGHVDVRSSMAAIRPHLQVC 886

841 POYNVLFDMLTVDHVMFYGRKLGSAVVGPEQDRLLQDVGLVSKOSVOTRHLSGQMOR 900
887 POYNVLFDMLTVDHVMFYGRKLGSAVVGPEQDRLLQDVGLVSKOSVOTRHLSGQMOR 946
901 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKYREGRTLLIISTHHLDEAEALLGD 960
947 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKYREGRTLLIISTHHLDEAEALLGD 1006
961 RVAVVAGRLCCCGSPFLRRHLGSGYLLVLVKARLPLTNEKADTMEGSDVTRQEKKN 1020
1007 RVAVVAGRLCCCGSPFLRRHLGSGYLLVLVKARLPLTNEKADTMEGSDVTRQEKKN 1066
1021 GSQGSRTGTPQLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTELA 1080
1067 GSQGSRTGTPQLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTELA 1126
1081 LRLTGYGISTDSEIEIFLKVVEECAADTDMEDSCGQHLCTGTAGLVDVTLRLKMPPOETA 1140
1127 LRLTGYGISTDSEIEIFLKVVEECAADTDMEDSCGQHLCTGTAGLVDVTLRLKMPPOETA 1186
1141 LENGEPAGSAPETDQSGPDAVGRVQGWALTROQLQALLKRFLLARRSRRLGFAQIVLP 1200
1187 LENGEPAGSAPETDQSGPDAVGRVQGWALTROQLQALLKRFLLARRSRRLGFAQIVLP 1246
1201 ALFVGLALVSLVPPFGHYPALELSPTMYGAQVSFESEADPGDPGRARLLEALLQAGL 1260
1247 ALFVGLALVSLVPPFGHYPALELSPTMYGAQVSFESEADPGDPGRARLLEALLQAGL 1306
1261 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNWTPSPSPACQCSOPGARLLPDCPAAAGGP 1320
1307 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNWTPSPSPACQCSOPGARLLPDCPAAAGGP 1366
1321 PPOAVTGSSEVVQNLFGRLNSDFLVKTYPRLVQGLKTKKVNVEVYGFSGFSGRDPGL 1380
1367 PPOAVTGSSEVVQNLFGRLNSDFLVKTYPRLVQGLKTKKVNVEVYGFSGFSGRDPGL 1426
1381 PSGOELGRSEELWALLSPLPGGALDRVLKNTAWAHSLSLDAQSLKTIWFKNGHWSVAF 1440
1427 PSGOELGRSEELWALLSPLPGGALDRVLKNTAWAHSLSLDAQSLKTIWFKNGHWSVAF 1486
1441 VNRSASNAI LRAHLPPGPARHAHSITTLNHPNLTKQESAAALMASSVDVLVSI CVVFAM 1500
1487 VNRSASNAI LRAHLPPGPARHAHSITTLNHPNLTKQESAAALMASSVDVLVSI CVVFAM 1546
1501 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWGNFLWDMCNLYVPACIVVLIFLAFQ 1560
1547 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWGNFLWDMCNLYVPACIVVLIFLAFQ 1606
1561 QRAVAPANLPALLLL LLLYVWSITPLMYPASPFPSVPSTAYVVLTCINLFIGINGSMAT 1620
1607 QRAVAPANLPALLLL LLLYVWSITPLMYPASPFPSVPSTAYVVLTCINLFIGINGSMAT 1666
1621 FVLEFSDQKLOEVSRIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1680
1667 FVLEFSDQKLOEVSRIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1726
1681 WEVVGNKLLAMVIOGPIFLPTLLQHRSQLLPQPRVRSPLPLIGEEDEVARERVRVQ 1740
1727 WEVVGNKLLAMVIOGPIFLPTLLQHRSQLLPQPRVRSPLPLIGEEDEVARERVRVQ 1786
1741 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPGCEFGLLGVNGAGKTSIFRMVYGTDLA 1800
1787 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPGCEFGLLGVNGAGKTSIFRMVYGTDLA 1844
1801 SRGSAVLAGHSVAREPSAAHLSMG-YCPQSDAIFELL 1836
1845 SFTHELLCIH-----HLLLTGYCM---PIFVLL 1869

RESULT 10

US-09-995-542-3

; Sequence 3, Application US/09995542

; Patent No. US20020127647A1

GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Iarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2121
; TYPE: PRT
; ORGANISM: Mus musculus
; ORGANISM: Mus musculus
US-09-995-542-3

Query Match 77.0%; Score 8389.5; DB 9; Length 2121;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;

QY 1 PPLEHECHFPNKPPLPSAGTVPMLQGLICNNVNTCFPQLTPGPEGRGLSNFNDLSVRL 60
DB 1 PPLEHECHFPNKPPLPSAGTVPMLQGLVCNNVNSCFQHTPGKPGVLSNFKDSLIRLL 60
QY 61 ADARTVLGGASAHRTLAGLKLATLRAARSTAP-----OPTKQSPLEPPMLDVAELLTS 116
DB 61 ADRTVLGGHSIODMLDALGKLPLVLRVAGGARGPQESDPTSQ-----SVTKLLEK 113
QY 117 LRRTESILGALGAQOAPLHSLLEAEADLAQELALRSLVELRALLORPGTSGPLELLE 176
DB 114 ILORASLDPLVQQAQSMKFSDAIRDLAQLLTLPSLMELRALLRRPGSAGSLVSE 173
QY 177 ALCSVRGSPSTVGSFNWYEASDLMELVQPEPSALPDSLSLSPACSELIGALDHPVRL 236
DB 174 ALCSVRGSPSTVGSFNWYEANQNEFMGEVAPALPDNSLSPACSEFVGTLDHPVRL 233
QY 237 LWRRLKPLILGKLFPADPTFTTKLMAQVNRTEETLLRDLREVWMLGPRIFTMDS 296
DB 234 LWRRLKPLILGKLFPADPTFTTKLMAQVNRTEETLLRDLREVWMLGPRIFTMDS 293
QY 297 SNVAMQLLOMDEGRORPRGGRDMEALRGLFDPGSGYSDQADAHVGLVGTGR 356
DB 294 TNVAMQLRLDVGCTGQRTQRTQRAKKLEAKDFLPSRGYSWRAHADMGRLAGLQ 353
QY 357 VTCELSLDKLEAAPSAAALVSRALQLLAEHRFWAGVVFGLGPESDSPTTEHTPDGLGPHV 416
DB 354 MMECVSLDKLEAVPSEALVSRALLEGERRLWAGIVFLSPHEPLDPSELSPALSPGHL 413
QY 417 RIKIRMDIDVVTNRKIRDFWDPGPAADPLTDLRYWVGGFVYLODLVEAAVRLSGAN 476
DB 414 RFKIRMDIDVVTNRKIRDFWDPGPAADPLTDLRYWVGGFVYLODLVEAAVRLSGAN 473
QY 477 PRAGLVLOQMPYPCYVDDVFLVLSLPLFLTLAMYSVTLTKAVRKEKTRLRUTMR 536
DB 474 SRTGLVLOQMPYPCYVDDVFLVLSLPLFLTLAMYSVTLTKAVRKEKTRLRUTMR 533
QY 537 AMGLSRAVLWGLWFLSCGLGPFLLSAAALLVLVLKLGIDILPYSHPGVWFLFLAAFAVATVQ 596
DB 534 AMGLSRAVLWGLWFLSCGLGPFLLSAAALLVLVLKLGIDILPYSHPGVWFLFLAAFAVATVQ 593
QY 597 SFLLSAPFSRANLAAACCGGLAYFSLYLPVLCVAMDRPLPAGGRVAASLSPVAFGCGE 656
DB 594 SFLLSAPFSRANLAAACCGGLAYFALYLPVLCVAMERLHLGGLLAASLSPVAFGCGE 653
QY 657 SLALLEEQGGAQWNVGTPTADVFSLAQVSGLLLDALYGLATWYLEAVCPGOYGP 716
DB 654 SLALLEEQGGAQWNVGTPTADVFSLAQVSGLLLDALYGLATWYLEAVCPGOYGP 713
QY 717 EPWNFPRRSYWCGRPPKSPAPCTPLDPKVLVEBAPGLSPGVSVRSLEKFPFGPQP 776
DB 714 EPWNFPRRSYWCGRPPKSPVLAQAPQDPKVLVEBAPGLSPGVSVRSLEKFPFGPQP 773


```
Qy 777 ALRGLSLDFYQHITAFIAGHNGAGKTTTILSILSGLFPSPGSAFIIIGHDVSRSSMAAIRPH 836
Db 774 ALQGLNLDFYEGHITAFIAGHNGAGKTTTILSILSGLFPSPGSAFIIIGHDVSRSSMAAIRPH 833
Qy 837 LQVCQYNVLFDMITVDHVVYFYGRLKGLSAAVGPBQDRLLQDVGLVSKOSVQTRHLSG 896
Db 834 LGICQYNVLFDMITVDHVVYFYGRLKGLSAAVGPBQDRLLQDVGLVSKOSVQTRHLSG 893
Qy 897 GQWRKLSVAIAFVGSSQVILDEPTAGVDPASRRGIGWELLILKYRGRITLILSTHLLDEAE 956
Db 894 GQWRKLSVAIAFVGSSQVILDEPTAGVDPASRRGIGWELLILKYRGRITLILSTHLLDEAE 953
Qy 957 LIGDRVAVVAGGRLLCCCGSPFLRRHLGSGYYILTVKARLPITTNE--KADTIDMEGSDVTR 1015
Db 954 LIGDRVAVVAGGRLLCCCGSPFLRRHLGSGYYILTVKSSQSLVTHDAKGDSE-----DPR 1008
Qy 1016 QKXNGSGS-----RVGTPQLLALVQHVWPGARLVEBEP 1050
Db 1009 REKSDGNGRTSDYATFRGTSKSNQAPAPGAVPITPSTARILELVQHVPGAQVLEDLP 1068
Qy 1051 HELVLVLPYTGADHGSFATLRELDTRLAELRLTGYSIDTSLSEIFLKVVEECAADTDM 1110
Db 1069 HELVLVLPYAGALDGSFANVQELDQQLLELGLTGYSIDNLEEIFLKVVED--AHREG 1126
Qy 1111 EDGSCGCHLCTGIAGLDVTLRLKMPPOBTALENGEPAGSA--PETDQSGDPDAGVRVQGW 1168
Db 1127 GDSRQLHLRT-----CTPQPTGPEASVLENGELAKLVLDPOAPQGLAPNA-AQVQGW 1179
Qy 1169 ALTROQLQALLKRPILLARRRGRGFAQIVLPALFVGLALVFSILVPPFGHYPALRLSPT 1228
Db 1180 TLTCQQLRALLKRPILLARRRGRGFAQIVLPALFVGLALVFSILVPPFGHYPALRLSPT 1239
Qy 1229 MYGAQVSFEADAGDCGRARLEALLOEAGLEPPVQHSR-----PSAPEV 1277
Db 1240 MYGPQVSFEADAGDCGRARLEALLOEAGLEPPVQHSR-----PSAPEV 1299
Qy 1278 PAEYAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTGSVEVQNL 1337
Db 1300 PPDVASILASGNWTPESPSPACQSQPGARRLLPDCPAGAGPPPPQAVAGLVEVQNL 1359
Qy 1338 GRNLSDFLVKTYPRIVROGLTKKWNVVRVGGFSLGDRDGLPGSGBELGRSVBELMALL 1397
Db 1360 GRNVSDFLVKTYPSLVRRLGLTKKWNVVRVGGFSLGDRDGLPGTGVVWRTLAIRALL 1419
Qy 1398 SPLPGCALDVLKNI-TAWAHSIDAOQSLKIWFNNKGWHSWAFVNRASNAILRAHLPGP 1457
Db 1420 SPQGNALDRILNNLTQWALGLDARNLSUKIWFNNKGWHSWAFVNRANGLLHALLPSGP 1479
Qy 1458 ARHAHSITTLNHPNLTKQELSEALMASVVDVLVSIQVWFAMGFVPASFTLVLEERTV 1517
Db 1480 VRHAHSITTLNHPNLTKQELSEALMASVVDVLVSIQVWFAMGFVPASFTLVLEERTV 1539
Qy 1518 RAKHLQMLGSLSPTYLWGNFLWDMCNVLPACIVVLIIFLAPQRAYVAPANLPALELLL 1577
Db 1540 RAKHLQMLGSLSPTYLWGNFLWDMCNVLPACIVVLIIFLAPQRAYVAPANLPALELLL 1599
Qy 1578 LLYGWSITPLMYPASFPFSPVSTAVVLTICNLFTIGINSMATFVLELFPDOKLQEVSR 1637
Db 1600 LLYGWSITPLMYPASFPFSPVSTAVVLTICNLFTIGINSMATFVLELFPDOKLQEVSR 1659
Qy 1638 LKQVFLIPFPFCLARGLDWVRNQAMADAFERLGRDQFQSPLRWEVVGKILLAMVIOG 1697
Db 1660 LKQVFLIPFPFCLARGLDWVRNQAMADAFERLGRDQFQSPLRWDIILGKILLAMVIOG 1719
Qy 1698 FLLFTLLQHRSLQPPQVRVSLPLGBEDBVARERVRVQGTQGGDVLVRLNLTQYR 1757
Db 1720 FLLFTLLQHRSLQPPQVRVSLPLGBEDBVARERVRVQGTQGGDVLVRLNLTQYR 1779
Qy 1758 GORPAPVDRICLIGPPGCGFLGVNGAGKSTFRMTVGTDLTASRGEAVLAGHSVAREPS 1817
Db 1780 GORPAPVDRICLIGPPGCGFLGVNGAGKSTFRMTVGTDLTASRGEAVLAGHSVAREPS 1839
```

```
Qy 1818 AAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVATAGSGLARLGLSWADRP 1877
Db 1840 AAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVATAGSGLARLGLSWADRP 1899
Qy 1878 GTYSGGNKRLATLALVGDPAVPLDPTTGMPSARRFLWNSLLAVVREGRSVNLTS 1937
Db 1900 GTYSGGNKRLATLALVGDPAVPLDPTTGMPSARRFLWNSLLAVVREGRSVNLTS 1959
Qy 1938 SMECEALCSRLAIAVNGRFRCLASPOHLKGRFAAGHTLTLRVPAARSOPAAAFVAA 1997
Db 1960 SMECEALCSRLAIAVNGRFRCLASPOHLKGRFAAGHTLTLRVPAARSOPAAAFVAA 2019
Qy 1998 GSELREAHGGRRLRQLPFGRCALARVEGELAVGAEGHGVDFSVQTMLEVFYFSGD 2057
Db 2020 GAELREVHGSRLRQLPFGRCALARVEGELAVGAEGHGVDFSVQTMLEVFYFSGD 2079
Qy 2058 QGKDBDTEQKEAGVGDVDPAPGLOHPKRVSOFLDDPSTAEVTL 2100
Db 2080 QGEBEESRQ-EABEEVEVKPGRQHPKRVSRFLEDPSSVETMI 2121

RESULT 11
US-09-995-542-2
; Sequence 2, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2167
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-995-542-2

Query Match 77.0%; Score 8389.5; DB 9; Length 2167;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;

Qy 1 PPLEHHECHFPNKPPLSAGTVPMWLOGLICNVNNTCFPQPTGPEBPGRLSNFNDLSVRL 60
Db 47 PPLEHHECHFPNKPPLSAGTVPMWLOGLICNVNNTCFPQPTGPEBPGRLSNFNDLSVRL 106
Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAO-----OPTKQSPLEPPMLDVNELLTS 116
Db 107 ADRTVLGGHSIQMDLALGKLIPLVRAVGGARQPSDQTSQ-----SVTKLLEK 159
Qy 117 LLRTESLGLALGQAQEPHLSLEAEADLAOELLALRSILVELRALLQRPRTSGPLELISE 176
Db 160 ILQASLDVPLGQAQDSKRFESDAIRDAQELLTLPMLRALLRRPRGSGSLELVSE 219
Qy 177 ALCSVRGSPSTVGSPLNMYEASDLMELVGOEPESALPDSSLSACSELIIGALDHPRL 236
Db 220 ALCSKTGFPSPGSLNMYEANQLNEFMGPEVAPALPDNSLSPACSEFVGTLDHPVRL 279
Qy 237 LWRELKPLILGLKLPADPTPTKRLMAQVNTFEELTLRRDVREVMELGRIETFMNDS 296
Db 280 LWRRLKPLILGLKLPADPTPTKRLMAQVNTFEELTLRRDVREVMELGRIETFMNDS 339
Qy 297 SNVAMLRLLQMQDEGRQRPRPGGRDHWEALRSFLDPGSGYQWQADAHADVGLVLTGR 356
Db 340 TNVAMLRLLQMQDEGRQRPRPGGRDHWEALRSFLDPGSGYQWQADAHADVGLVLTGR 399
Qy 357 VTECLSLDKLEAPSEAAVLSRALQALAEHRFWAGVFLGPEPDSDDPTEHTPTDLPQGHV 416
Db 357 VTECLSLDKLEAPSEAAVLSRALQALAEHRFWAGVFLGPEPDSDDPTEHTPTDLPQGHV 416
```


Db 400 MMECVSLDKLEAVPSEBVALSRALLGERLWAGIVFLSPHPLDSELSSPALSPCHL 459
Qy 417 RIKIRMDIVVTRTKIRDRFWDGPAADPTDURYVWGGFVYLQDLVERAAVVLGSAN 476
Db 460 RFKIRMDIDVTRTKIRDKFWDGPAADPMDLRYVWGGFVYLQDLLEQAAVVLGGCN 519
Qy 477 PRAGLYLQOMPYPVCVDDVFLRVLRSRIPLPLTLTAWIYSVTLTKVAVREKETRLDTMR 536
Db 520 SRTGLYLQOMPHPCCVDDVFLRVLRSRLPLTLTAWIYSVALTKVAVREKETRLDTMR 579
Qy 537 AMGLSRAVLWLGWFLSCLGPPFLSAAALLVLVLKGLDILPYSHPGWFLFLAAFAVAVTQ 596
Db 580 AMGLSRAVLWLGWFLSCLGPPFLSAAALLVLVLKGNILPYSHPVVIFLFLAAFAVAIAQ 639
Qy 597 SFLLSAFSFRANLAAACGLAYFSLYPLVLCVAVWRDLRPAAGRVAAASLLSPVAFGCE 656
Db 640 SFLLSAFSFRANLAAACGLAYFALYPLVLCVAVWRDLRHLGGLLAAASLLSPVAFGCE 699
Qy 657 SLALLEQGEAGAHNVCTRTADVFSIAOVSGILLDAALYGLATWYLEAVCPGOYGP 716
Db 700 SLALLEQGEAGAHNVCTRTADVFSIAOVSGILLDAALYGLATWYLEAVCPGOYGP 759
Qy 717 EPMNFPFRSRYWCGPPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSRSLKRPSPGP 776
Db 760 EPMNFPFRSRYWCGPPPKSSVLAPAPQDPKVLVEEAPPGLSPGVSRSLKRPSPGP 819
Qy 777 ALRGLSLDFYOGHTAFIHLNGAGKTTLSILSGLFPSPGSGSAPILGHVSRSSMAAIRPH 836
Db 820 ALQGLNLDFYEGHTAFIHLNGAGKTTLSILSGLFPSPGSGSAPILGHVSRSSMAAIRPH 879
Qy 837 LGVCPQNVNLFDMTLVDEHVWFYGRGLGSAVVGPEQDRLQDVGIVSKSVQVTRHLSG 896
Db 880 LGICPQNVNLFDMTLVDEHVWFYGRGLGSAVVGPEQDRLQDVGIVSKSVQVTRHLSG 939
Qy 897 GMQRKLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAE 956
Db 940 GMQRKLSVAIAFVGGSRVIMDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAE 999
Qy 957 LLGDRVAVAGRLCCGSPFLFRHLGSGYLLTLVKARLPLTNE-KAOTDMEGSVDR 1015
Db 1000 LLGDRVAVAGRLCCGSPFLFRHLGSGYLLTLVKRSQSLVTHDAKGDSE-----DPR 1054
Qy 1016 QEKKNQSGS-----RVGTPLLALVQHVPGARLVEELP 1050
Db 1055 REKSDNGRTSDTAFRTGTSKSNQAPAGVPIPTSTARIELVQHVPGARLVEELP 1114
Qy 1051 HELVLVLYTCAHDSFATLFRLELRLABELRTGYGSDTSLSEEIFLKVVECAADTM 1110
Db 1115 HELVLVLYAGALDGSFAMVFOELDQQLLELGLTGYGSDTNLEEIFLKVVED-AHREG 1172
Qy 1111 EDGSCGHLCTGAGLDVTLRKWPPQETALENGEPAGSA--PETDQSGPDAVGRVQGW 1168
Db 1173 GDSRPQLHLRT-----CTPQPTGPEASVLENGELAKLVLDPOAQGLAPNA-AQVQGW 1225
Qy 1169 ALTRQOQLALLKRLFLARSRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1228
Db 1226 TLTCQQRALLKRLFLARSRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1285
Qy 1229 MYGAQVFFSEADPGDFGRARLLLEALQEALEPPVQHSR-----FSAPEV 1277
Db 1286 MYGQVFFSEADPGDFGRARLLLEALQEALEPPVQHSR-----FSAPEV 1345
Qy 1278 PAEVAKVLASGNWTPESPSPACQSQGARRLLPDCPAAAGGPPPOAVTGSSEVONLT 1337
Db 1346 PPDVASTLASGNWTPESPSPACQSQGARRLLPDCPAGAGGPPPOAVTGSSEVONLT 1405
Qy 1338 GRNLSDFLVKTYPRVRLQGLTKKVVNEVRYGFSGLGRDPGLPSGOELGRSVEELWALL 1397
Db 1406 GRNVSDFLVKTYSRLVRGLTKKVVNEVRYGFSGLGRDPGLPSGOELGRSVEELWALL 1465
Qy 1398 SPIPGGALDRLVKNLTAWAHSLLDAQDSKLTWNNKNGHSMVAFVNRASNAIRLAHLPGP 1457
Db 1466 SPOGNAIDRLNLTQWALGLDARNSLKTFWNNKNGHSMVAFVNRANGLHALLPSGP 1525

Qy 1458 ARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVPFAMSFVPASPFTLVLIERVT 1517
Db 1526 VRHAHSITTLNHPNLNLTKEQLSEATLIASSVDVLVSIQVVPFAMSFVPASPFTLVLIERIT 1585
Qy 1518 RAKHLQALMGGLSPTLWYGNFLWDMCNVLPACIVVLIFLAFOORA VVAPANLPALALLL 1577
Db 1586 RAKHLQALVSLGPOTLWYGNFLWDMCNVLPACIVVLIFLAFOORA VVAPANLPALALLL 1645
Qy 1578 LLYXGWSITPLMYSPASFFSVPSTAYVVLITCINLFIGINGSMAFVLEFSDQKLQEVRSI 1637
Db 1646 LLYXGWSITPLMYSPASFFSVPSTAYVVLITCINLFIGINGSMAFVLEFSDQKLQEVRSI 1705
Qy 1638 LKQVFLIPFHFCLGRGLIDMVRNQAMADAPERIGDRQFQSPQLRWEVVGKLLAMVIOGPL 1697
Db 1706 LKQVFLIPFHFCLGRGLIDMVRNQAMADAPERIGDRQFQSPQLRWEVVGKLLAMVIOGPL 1765
Qy 1698 FLIITLLOHRSOLLQPPRVRSIPLIGEEDVABERERVOGATOGDVLVLRNLTKVYR 1757
Db 1766 FLIITLLOHRSOLLQPPRVRSIPLIGEEDVABERERVOGATOGDVLVLRNLTKVYR 1825
Qy 1758 GQMPAVDRICLGIIPGECFCGLLGVNGAGKTSIFRMVGTDLASRGEAVLAGHSVAREPS 1817
Db 1826 GQMPAVDRICLGIIPGECFCGLLGVNGAGKTSIFRMVGTDLASRGEAVLAGHSVAREPS 1885
Qy 1818 AAHLHWYCPQSDAIFELTGTREHLELLARLGVPEAQVAQTAGSGLARLGLSWYADRP 1877
Db 1886 AAHLHWYCPQSDAIFELTGTREHLELLARLGVPEAQVAQTAGSGLARLGLSWYADRP 1945
Qy 1878 GTYSGGNKRLATALAVGDPVAVFLDEPTTGMDSARRFLMNSLLAVVREGSVMLTSH 1937
Db 1946 GTYSGGNKRLATALAVGDPVAVFLDEPTTGMDSARRFLMNSLLAVVREGSVMLTSH 2005
Qy 1938 SMEECALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTLFLRVPAAARSQPAAPAAAEFP 1997
Db 2006 SMEECALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTLFLRVPAAARSQPAAPAAAEFP 2065
Qy 1998 GSELREAHGRRLRFQPLPGRCALARVFGELAVGAEHGVEDFSVQTMLEEVFLYPSKD 2057
Db 2066 GSELREAHGRRLRFQPLPGRCALARVFGELAVGAEHGVEDFSVQTMLEEVFLYPSKD 2125
Qy 2058 QGKDEDETEQKAGVGVDPAPLQHPKRVSOPLDDPSTAEVTL 2100
Db 2126 QGEESSRQ-EABEEVSKPQHPKRVSRFLEDPSVETMI 2167

RESULT 12

US-09-995-542-8
; Sequence 8, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1550
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-8

Query Match 69.7%; Score 7596; DB 9; Length 1550;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 121 ESIGLALGOAQEPHLSLEAAEDLAQELLARSLVELRALLQRPCTSGPFLLEALCS 180

Db 29 ESIGLALGOAQLSHSLAEAGDLAQELLALSLVELRALLQKRGTSGLLEALLCS 88
Qy 181 VEGPSTVCPSPNWNVEASDLMEVGOEPESALPDSSLSPACSELIGALDHPHLLWRR 240
Db 89 VEGPSTVCPSPNWNVEASDLMEVGOEPESALPDSSLSPACSELIGALDHPHLLWRR 148
Qy 241 LKPLILGLKLLFAPDTPFTTKLMAQVNRTEPEELTLDRDREVWEMLGPRITFTPMNDSSNVA 300
Db 149 LKPLILGLKLLFAPDTPFTTKLMAQVNRTEPEELTLDRDREVWEMLGPRITFTPMNDSSNVA 208
Qy 301 MLQRLQWODEGRORPRPGGRDHMEALRSFLDPGGGYSWQDAHADVGHVLTGLGRVTEC 360
Db 209 MLQRLQWODEGRORPRPGGRDHMEALRSFLDPGGGYSWQDAHADVGHVLTGLGRVTEC 268
Qy 361 LSLDKLEAPSEAAALVSRAQLLALHREFWAGVVFILGPEDSSDPTHEPTDPLGPHVRIKI 420
Db 269 LSLDKLEAPSEAAALVSRAQLLALHREFWAGVVFILGPEDSSDPTHEPTDPLGPHVRIKI 328
Qy 421 RMDIDVTRTKIRDFWDPGPAADPLTDLRYVMGFFVYLODLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRDFWDPGPAADPLTDLRYVMGFFVYLODLVERAAVRVLSGANPRAG 388
Qy 481 LYLOQMPYPCYVDDVFLRVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 540
Db 389 LYLOQMPYPCYVDDVFLRVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 448
Qy 541 SRAVLWGLWFLSCIGPFLLSAALLVVLKGLDILPYSHPGVVFLLAFAFVATVTSFLL 600
Db 449 SRAVLWGLWFLSCIGPFLLSAALLVVLKGLDILPYSHPGVVFLLAFAFVATVTSFLL 508
Qy 601 SAFFSRANLAAACGLAFSLYLPVLCVANDRLPAGRVAASLLSPVAFGFCESIAL 660
Db 509 SAFFSRANLAAACGLAFSLYLPVLCVANDRLPAGRVAASLLSPVAFGFCESIAL 568
Qy 661 LEEQEGQAHNVGTRPTADVFLSAQVSGLLDALLDAALYGLATWYLEAVCPQYGIPEPN 720
Db 569 LEEQEGQAHNVGTRPTADVFLSAQVSGLLDALLDAALYGLATWYLEAVCPQYGIPEPN 628
Qy 721 PFRRSYWCGRPPKSPAPCTPLDPKVLVEAPPLGSPGVSVRSLEKRPFGSPOPALRG 780
Db 629 PFRRSYWCGRPPKSPAPCTPLDPKVLVEAPPLGSPGVSVRSLEKRPFGSPOPALRG 688
Qy 781 LSLDFYQCHIATFLHNGAGKTTLSILSGLFPSPGSAFILGHVDRSSMAAIRPHLVC 840
Db 689 LSLDFYQCHIATFLHNGAGKTTLSILSGLFPSPGSAFILGHVDRSSMAAIRPHLVC 748
Qy 841 POYNVLFDMLTVDHWFYGRKGLSAAVGPEDRLLODVLVSKQSVQTRHLSGGMQR 900
Db 749 POYNVLFDMLTVDHWFYGRKGLSAAVGPEDRLLODVLVSKQSVQTRHLSGGMQR 808
Qy 901 KLSVAIAFVGSQVILDEPTAGVDPPASRGWELLKLYREGRTILSTHLDLAEALLGD 960
Db 809 KLSVAIAFVGSQVILDEPTAGVDPPASRGWELLKLYREGRTILSTHLDLAEALLGD 868
Qy 961 RVAVAGRLCCGSPFLRRHLSGGYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
Db 869 RVAVAGRLCCGSPFLRRHLSGGYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN 928
Qy 1021 GSQSRVGTQPLLALVQHVPFCARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAE 1080
Db 929 GSQSRVGTQPLLALVQHVPFCARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAE 988
Qy 1081 LRLTYGIGSDTSLBIEIFUKVVEECAAUPTMEDGCGQHLCGTGAGLDVTLRLKMPQOSTA 1140
Db 989 LRLTYGIGSDTSLBIEIFUKVVEECAAUPTMEDGCGQHLCGTGAGLDVTLRLKMPQOSTA 1048
Qy 1141 LENCEPAGSAPETDQSGPDVAVRQGWALTRQOQLALLKRFLLARESGLEFAQIVLP 1200
Db 1049 LENCEPAGSAPETDQSGPDVAVRQGWALTRQOQLALLKRFLLARESGLEFAQIVLP 1108
Qy 1201 ALFVGLALVFLSVLPPFGHYPALRLSPMTYGAQVSFFSEDAFGDPGRARLLLEALLQEAGL 1260

Db 1109 ALFVGLALVFLSVLPPFGHYPALRLSPMTYGAQVSFFSEDAFGDPGRARLLLEALLQEAGL 1168
Qy 1261 BEPPVQSHSHRPSADEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1320
Db 1169 BEPPVQSHSHRPSADEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1228
Qy 1321 PPPQAVTSGGEVQVQNLGTGRNLSDFLVKTYPRLVROGLTKTKWVNEVRYGGFSLGGDRDPL 1380
Db 1229 PPPQAVTSGGEVQVQNLGTGRNLSDFLVKTYPRLVROGLTKTKWVNEVRYGGFSLGGDRDPL 1288
Qy 1381 PSGQLGRSVBELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGWMSWAF 1440
Db 1289 PSGQLGRSVBELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGWMSWAF 1348
Qy 1441 VNRSNATLRAHLPPGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLVVICVVFAM 1500
Db 1349 VNRSNATLRAHLPPGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLVVICVVFAM 1408
Qy 1501 SFVPASFTVLIEERVTRAKHLQLMGGLSPTLYMLGNFLWDMCNLYVPACIIVLIFLAQ 1560
Db 1409 SFVPASFTVLIEERVTRAKHLQLMGGLSPTLYMLGNFLWDMCNLYVPACIIVLIFLAQ 1468
Qy 1561 QRAYVAPANLPAALLLLLYGWSITPL 1587
Db 1469 QRAYVAPANLPAALLLLLYGQAVPV 1495

RESULT 13

US-10-313-641-9
; Sequence 9, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P02351US2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-313-641-9

Query Match 51.5%; Score 5610.5; DB 14; Length 2261;

Best Local Similarity 50.0%; Pred. No. 0;

Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;

Qy 1 PPLEHHECHFPNKPSPAGTVPMVLOGLICNVNNTCFPOLTPGEBPGRLSNPNDSLSRLL 60

Db 47 PPYEHECHFPNKPSPAGTVPMVLOGLICNVNNTCFPOLTPGEBPGRLSNPNDSLSRLL 106

Qy 61 ADARTVLGGASAHRTLAGLKLIALTAARSTAQ-----PQ 96

Db 107 SDARLLLYSQKDTSMKMRKVLRTLOQIKKSSNLKLODFLVNDETFSGLYHNLSLPK 166

Qy 97 PTQSPLEPPML-----DVAEL----- 113

Db 167 STVDKMLRADVLHKVFLQYQLHTSLCNGSKSEMIQLGDQEVSELGIPREKLAAR 226

Qy 114 -----LTSLLRT--ESLGLALGOAQBPFLSHLEAAEDLAQELLALRSVLBEL-- 159

Db 227 RVLRSNMDILKPIRLTINSTSPPSKELABATKTLHLSGLTLAGELFSMRSDMRQV 286

Query Match	51.5%	Score	5610.5	DB	14	Length	2261
Best Local Similarity	50.0%	Pred.	No. 0				
Matches	1118	Conservative	361	Mismatches	596	Indels	161
Gaps	26						
QY	1	PPLEHCHCPNKP	PLPSAGTVPMLQGLICNNVNTCFPQLTPGEPGRGLSNFNDLSVLRLL	60			
Db	47	PPYEQHCHCPNKA	PSAGTLPWQGLICNANPCFRYPYPGEPAGVGVGNFNSIVARLF	106			
QY	61	ADARTVLGGSAS	AHRTLGLAGKLIATLRAASTAQ-----PQ	96			
Db	107	SDARRLLYSQ	KDTSMDKMRKVLTLLQIKKSSSNLKQDPLVDNETFSGFLYHNLSLPK	166			
QY	97	PTQSPLEPPM	L-----DVABL-----	113			
Db	167	STVDMRLRAD	VILHKVFLQYQLHLTSLCNGSKSEEMIQLDQDQVSELCLGPREFKLA	226			
QY	114	-----LTS	LRT--ESLGLALGAQAPLHSLLEAAEDLAQELALASLSVELR--A	159			
Db	227	RVLRSNMILK	PIRLTUNSTSPPSKELAAATKLLHSLGLTQAELEFSMSWSMDRQEV	286			
QY	160	LLQRPGTSG	PLEL---LSBALCSVRGPSSTVGPSLWNYEASDLMELVG---QEPESAL	212			
Db	287	FLTNVNSSS	STQIYQAVSRIVCPHPPGGGLIKSLNWYEDNNYKALFGGNGTEEDAET	346			
QY	213	PDSSLSPAC	SICALDHPRLSLLRRLKPLILGLKLLFAPDPTFKLMAQVNRTEEL	272			
Db	347	YDNSTTPY	CNDLMKNLESSPLSRIIKALKPLVKGILYTPDTPATQVMAEVNKTQEL	406			
QY	273	TLLRDVRV	WEMLGPRIFTFMNDSSNVAMQLRLIQMDEGR-ROPRPGGRD-----	322			
Db	407	AVFHDLEG	MWELSPKIWTFMENSQEMDLVRMLLDSRDNDHPFEQQLDGLDWTQA	466			
QY	323	---HMEALR	SFLDPGSGG---YSQDADHADVGHVLVGTUGRVTECLSLDKLSAAAPSEAA	377			
Db	467	LAKHPED	VQS-----SNGSVYTVREAFNETNQAIRTISRPEMVCNMLKLEPIATE	521			
QY	378	RALQLLA	EHFAGVWFLGPEDDSDPTHEPTDPLGPGHVRIKIRMDIDVYTRTNKIRDR	437			
Db	522	KSMELDER	KFWAGIVFTGTPGSIELPH-----HVYKIRMDIDNVERTNKIKDGY	573			
QY	438	WDGPAAD	PLTDLRYVWGGFYVLQDLVERAAVRVLGSANPRAGLYLQOMPYPYCVDD	497			
Db	574	WDGPRAD	PEDMRVYWGGFAYLDQVVEQAIIRVLGTGTEKTKGVYMQMPYPYCVDD	633			
QY	498	RVLRSLS	PLFLTLAWIYISVTLTKAVVREKETRLDTRMRANGLSRAVLMLGWFLSC	557			
Db	634	RVMRSN	PLFMTLAWIYISVAIIKGIYVEKARUKETMRIMGLDNSLWFWSWISS	693			
QY	558	LLSAALL	VLVLKGLDILPYSPHGVVFLFLAAFAVATVTSQFLLSAFSPSRANLAAAC	617			
Db	694	LVSAGLL	VLVLKGNLLPYSDPSVVFVLSVFAVVTILQCLFLSTLPSRANLAAAC	753			
QY	618	YFSLYL	YPYVLVCVWRDRKLPAGGRVAASLLSPVAFGFCESLALLEBOGEQAOHNV	677			
Db	754	YFTLYL	YPYVLVCVWQDYVGFETKIFASLSPVAFGFCGYFALFEEGIGVQDWNLF	813			

Db 1993 EDVRRQRILDDGGQNDILEIKELTKIYRRKRPVDRICVGLPPGECFGLGVNGAGK 1952
Qy 1788 TSTFRMTGDTLARGSEAVLAGHVSAREPSAAHLSMGYCQSDAIFELLTGRHLELLAR 1847
Db 1953 SSIFKMLTGTDTTRGDAFLNKSILSIHIEHVQNGYCYQCFDAITELLTGRHVEFFAL 2012
Qy 1848 LRGPVPAQVAGTAGGLARGLSWYADRPAGTYGGNKRKLATATALVGDPAVVFLDEPT 1907
Db 2013 LRGPVEKGVKGWEAIRKGLVKYGEKAGYAGNTSGGNKRKLSTAMALIGGPPVVFLDEPT 2072
Qy 1908 TGMDDPSARRFLWNSLAVVREGSVMLTSHSMECEALCSRLAIVMNGRCFLGSPHLLK 1967
Db 2073 TGMDDPKARRFLWNCALSVKREGSVMLTSHSMECEALCTRMAIVMNGRCFLGSPHLLK 2132
Qy 1968 GRFAAGHTLRLVPAARS--QPAAFVAABFPGSELREAHGGRRLRFOLPPGRCALARVF 2025
Db 2133 NREGDGYTIVRIAGSNPDLKPQODFFGLAPPGSVLKEKHRNMLQYL--PSSLSSLARIF 2191
Qy 2026 GELAVHGAEGHVEDFVSQTMLEBEVLYFSKQDKDB---DTBEQKEAGVGVDPAPGLQH 2082
Db 2192 SILSQSKRLHIEDYSVSQTTLDQVFNFAKQSDDDHLLKDLHLKQTV--VDVAV---- 2246
Qy 2083 PKRVSQFLDDPSTAET 2098
Db 2247 ---LTSFLQDEKVKES 2259

RESULT 15
US-10-428-551-9
; Sequence 9, Application US/10428551
; Publication No. US20030229062A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P0231US3
; CURRENT APPLICATION NUMBER: US/10/428,551
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-428-551-9

Query Match 51.58; Score 5610.5; DB 14; Length 2261;
Best Local Similarity 50.08; Pred. No. 0;
Matches 118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
Qy 1 PPLEHHECHFPNKPPLPSAGTVPMLQGLICNNVNTCPQLTPGEBPGRLSNFDNLSVRL 60
Db 47 PPVEQHECHFPNKAAPSAGTLPWVQGLICNANNPCFRYPPTGAPGVVGNFNKSIARLF 106
Qy 61 ADARTVLGGSASHTLAGLKLATLRAASTAQ-----PQ 96
Db 107 SDARLLLYSQKOTSMDKMRKVLRTLOQIKSSSNLKLQDFLDVNETFFGLYHNLSPK 166
Qy 97 PTQSPLEPPML-----DVAEL----- 113
Db 167 STVDKMLRADVILHKVFLQYQLHLTSLNKGSKSEEMIQGDQEVSELCLGPKEKLAEE 226
Qy 114 -----LTSILRT--ESLGLAQOEPFLSHLEAAEDLAQELALASLVELR---A 159
Db 227 RVLRSNMDILKPIILRTLNSTSPFSPKELAEATKLLHSLGLTLAGELFMSMSWSDMRQEV 286

Qy 160 LLQRPRTGTSGLBL---LSEALCSVRGSPSSTVGPISLWNYEASDLMELVG---QEPESAL 212
Db 287 FLTNVSSSSSTQYQAVSRIVCGHPGEGGLKIKSLNWYEDNNYKALFCNGTEEDAETP 346
Qy 213 PDSLSLSPACSELGALDSHPLSRLWRLKPLILGKLLFAPDPTFTFKLMAQVNRTEEL 272
Db 347 YDNSTTPYCNLDLMKNLESSPLSRIIWKALKPLVLGKILYTPDTPATQVMAEVNKTPOEL 406
Qy 273 TLLRDVREVMELGPRIFTFMDNDSNVAMLQRLQLQWDEGR--RQPRPGGRD----- 322
Db 407 AVFHDLEGWMEELSPKITWTFMENSQEMDLVRMLDSRDNDHFWEQQLDGLDWTQAQDIVAF 466
Qy 323 ---HMEALRSFLDPGSGG--YSWQDAHADVGHVLTGRVTECLSDKLXAAASEAALVS 377
Db 467 LAKHPEDVQS-----SNGSVYTWREAFNETNQAIIRISFMECVNLNKLPIATEVWLN 521
Qy 378 RALQLLAEHRFWAGVVLGPEDESDPTHEPTDPLGPHVRIKIRMDIDVTRNKIRDRF 437
Db 522 KSMELDDERKFWAGIVFTGTPGSIELPH-----HVYKIRMDIDNVERTNKIKDGY 573
Qy 438 WDPGPAADPLTDLRYVMGPFVYLQDLVERAAVRLSGANPRAGLYLQOMPYPCYVDVFL 497
Db 574 WDPGPRADPEDRYVMGPFAYLQDVVEQAIIRVLTGTEKTKGVYMQOMPYPCYVDIFL 633
Qy 498 RVLRSPLPLTLAWIYSVTLVKAVVRKETEKLRTDMRAMLGRSLRAVLWGLFSLCLGPF 557
Db 634 RVMSRSMPLMTLAWIYSVAVIKIGIYVEKEARLKETMRIMGLDNLFWSPFISSLIPL 693
Qy 558 LLSAALLVLVLKGLDILPYSHPGVFLFLAAFAVAVTQTSFLLSAFPSRANLAAAGCLA 617
Db 694 LVSAGLLVLKUGNLLPYSDPSVVFVLSVFAVVTILQCLFLISTLFSRANLAAACGGII 753
Qy 618 YFSLYLPYVLCVAWRDLRPAGRVAAASLLSPVAFGFCESLALLEQEGEQAQHNVTGP 677
Db 754 YFTLYLPYVLCVAWQDVGVGFTLKIIFASLLSPVAFGFCGYFALFEEGICGVQNDLFPSP 813
Qy 678 T-ADVFSLAQVSGLLLDAAALYGLATWYLEAVCPQGYQGYPEPWNFFPRRYSWGPSPKS 736
Db 814 VEEDGFLNLTSSVSMMLFDTFLYGVMWTYIEAVFPQYQYIPRPWYFPCTKSYWFEESDEK 873
Qy 737 PAPCTPLD--PKVLVEEAPPLSPGVSRSLERFPGPSOPALRGLSLDYPQCHITAFGL 795
Db 874 SHPGSNQKRISIECMEEBPHLKLGVSIQNLVYVDGKMKVAVDGLALNIEGQITSFLG 933
Qy 796 HNGAGKTTTLLSILGLFPSPGSAFILGHVDVRSMAAIRPHLGVCPQYNYVLFMLTVDHEH 855
Db 934 HNGAGKTTTMSILTLFPPTSGTAYILGKDIRSEMSTIRQNLGVCVQHNVLFDMLTVEEH 993
Qy 856 WPFYGRKLGLSAAVGPBQDRLLQDVGL--VSKQSVQTRHLSGMGQRKLSVAIAFVGSQV 914
Db 994 IWFYARLKLGLSEKHVKAEMEQAQMDVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGSKV 1053
Qy 915 VILDEPTAGVDPASRRGIWELLILKYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 974
Db 1054 VILDEPTAGVDVPSRRGIWELLILKYRQRTIILSTHMDADVIGDRILAIISHGLCCVG 1113
Qy 975 SPLFLRRHLGSGYYLLTLVKARLPLTTN-----EKADTDMEGSVDTTQEKNGS 1022
Db 1114 SSFLXNLQGLTGYLLTLVKDVESSLSSCNSSSTVSYLKKEKDSVSSQSSDAGLSGSHES 1173
Qy 1023 QGSRVGTPLLALVQHWVPGARLVBELPHLVLVLPYTGADHSGSFATLRELDTRLAELR 1082
Db 1174 DTLTIIDVSAISNIRKHSPEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRISDLG 1233
Qy 1083 LTCVIGSDTSLBEIFLKVVEECAADTDMEDSGQHLCTGAGLDVTLRLKMPQETALE 1142
Db 1234 ISSYGLSETTLEBIFLKVAEESGVDAETSDGTLPARRNRAGF--DKQSCLRPTTEDAAD 1292
Qy 1143 --NGEPAQSAPETDQSGPDVAG--RVQGWALTRQQLQALLKRFLLARRSRGLFAQIV 1198
Db 1293 PNDSDDIPESRETDLISGMDGKGSYQVKGWKLTOQQFVALLWKRLIARRSRKGFQIV 1352

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 177.061 Seconds
(without alignments)
4254.650 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHHEFFNKPPLPSAGT.....QHPKRVSQLDDPSTAETVL 2100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 359729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10896	100.0	2146	6	Abg72696 Human ATP
2	10883	99.9	2146	4	Aau04483 Human PD-
3	10883	99.9	2180	5	Aaol14210 Human tra
4	10881	99.9	2146	5	Abp52096 Homo sapi
5	10881	99.9	2146	6	Abu08466 Human ABC
6	10880	99.9	2144	5	Aau09174 Human tra
7	10880	99.9	2144	7	Add37429 Human tra
8	10392.5	95.4	2059	6	Abu54629 Human NOV
9	10264	94.2	2008	6	Abu08464 Amino aci
10	10171.5	93.4	1993	6	Abu08465 Amino aci
11	9213.5	84.6	1873	4	Aau04484 Human PD-
12	8389.5	77.0	2167	6	Abg72695 Mouse ATP
13	7596	69.7	1550	6	Abg72697 Human ATP
14	5611.5	51.5	2261	3	Aab38111 Human ABC
15	5611.5	51.5	2261	3	Aab38114 Human ABC
16	5610.5	51.5	2261	3	Aab38117 Human ABC
17	5610.5	51.5	2261	3	Aab38115 Human ABC
18	5610.5	51.5	2261	3	Aab38109 Human ABC
19	5610.5	51.5	2261	3	Aab38082 Human ABC
20	5610.5	51.5	2261	3	Aab38112 Human ABC
21	5610.5	51.5	2261	4	Aab71749 Human ABC
22	5610.5	51.5	2261	4	Aab31361 Amino aci
23	5610.5	51.5	2261	4	Aab31365 Amino aci
24	5610.5	51.5	2261	6	Abu11899 Human ATP
25	5610.5	51.5	2261	6	Abt62034 Human ABC

26	5610.5	51.5	2261	6	ABR62033 Human ABC
27	5610.5	51.5	2261	7	ADP65173 Human ATP
28	5609.5	51.5	2261	3	Aab38116 Human ABC
29	5608.5	51.5	2261	3	AAB38113 Human ABC
30	5606.5	51.5	2261	3	AAB38110 Human ABC
31	5606.5	51.5	2261	3	AAB38105 Human ABC
32	5606.5	51.5	2261	4	AAB31362 Amino aci
33	5606.5	51.5	2261	4	AAB31366 Amino aci
34	5602.5	51.4	2261	4	AAB31363 Amino aci
35	5602.5	51.4	2261	4	AAB31367 Amino aci
36	5600	51.4	2260	3	AAB38106 Human ABC
37	5598.5	51.4	2261	3	AAB38104 Human ABC
38	5594.5	51.3	2261	5	ABB83120 Polymorph
39	5594.5	51.3	2261	5	ABB83116 Polymorph
40	5594.5	51.3	2261	5	ABB83124 Polymorph
41	5594.5	51.3	2261	5	ABB83122 Polymorph
42	5593.5	51.3	2259	3	AAB38107 Human ABC
43	5593.5	51.3	2261	4	AAM50228 Human ATP
44	5593.5	51.3	2261	4	AAE13022 Human ATP
45	5593.5	51.3	2261	4	AAU02176 Human ABC

ALIGNMENTS

RESULT 1
ABG72696
ID ABG72696 standard; protein; 2146 AA.
XX AC ABG72696;
XX DT 10-MAR-2003 (first entry)
XX DE Human ATP-binding cassette transporter-like protein, ABCL.
XX KW Human; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypohalimus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..46
Protein	/label= Signal_peptide 47..2146 /label= Mature_ABCL
US2002127647-A1.	
12-SEP-2002.	
28-NOV-2001; 2001US-00995542.	
28-NOV-2000; 2000US-0253520P.	
(SHUT//) SHUTTER J.	
(ULIA//) ULIAS L.	
Shutter J, Ulia L;	
WPI; 2003-147394/14.	
N-PSDB; ABX14666.	

Novel ATP-binding cassette transporter-like polypeptides and polynucleotides useful for diagnosing, preventing, treating disorders

PT involving immune, nervous system, thyroid, hypothalamus and impaired
PT transport of lipids.

XX Claim 13; Fig 2; 149pp; English.

XX The invention relates to an isolated murine and human ATP-binding
CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
CC encoded by the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110 or PTA-
CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
CC vectors, host cells, ABCL binding agents, a selective binding agent or
CC its fragment comprising at least one complementarity determining region
CC (CDR) with specificity for ABCL which (produced by immunising an animal
CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
CC fusion polypeptide, a device comprising a membrane suitable for
CC implantation (permeable to the protein and impermeable to materials
CC detrimental to the cells, and cells encapsulated within the membrane)
CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
CC and modulators are useful for the diagnosis and/or treatment of diseases
CC and conditions involving impaired transport of lipids, including
CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
CC involving functional and trophic disturbances of the nervous system such
CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
CC fibrosis, conditions involving multidrug resistance, conditions involving
CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
CC neutropenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies, autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL
XX

SQ Sequence 2146 AA;

Query Match 100.0%; Score 10896; DB 6; Length 2146;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 PPLEHHECHFPKPLPSAGTVPMVQGLICNVNNTCFPLTPGEPGRSLNFDNLSRL 60
47 PPLEHHECHFPKPLPSAGTVPMVQGLICNVNNTCFPLTPGEPGRSLNFDNLSRL 106
61 ADARTVLGGASHTLGLKGLIATLRAASTAQPTKOSPLPPMLDVAELTSLRT 120
107 ADARTVLGGASHTLGLKGLIATLRAASTAQPTKOSPLPPMLDVAELTSLRT 165
121 ESLGLAQOQEPPLHSLLEAEADLAQELLALRSIVELRALLOPRGTSGLPLELSEALCS 180
167 ESLGLAQOQEPPLHSLLEAEADLAQELLALRSIVELRALLOPRGTSGLPLELSEALCS 226
181 VRGSSSTVPSLNWYEASDLMEYVQGPESALPOSSLSPPACSELIGALDHPHLSRLWRR 240
227 VRGSSSTVPSLNWYEASDLMEYVQGPESALPOSSLSPPACSELIGALDHPHLSRLWRR 286
241 LKPLILGKLLFAPPTPFRKLMAQVNTFEBTLRLDRVREVMELGPRIETFMNDNSVA 300
287 LKPLILGKLLFAPPTPFRKLMAQVNTFEBTLRLDRVREVMELGPRIETFMNDNSVA 346
301 MLQRLQWODGRRQPRGGDHMEALRSFLDPGSGGYSWQDAHADYGHVGLTGRVTEC 360
347 MLQRLQWODGRRQPRGGDHMEALRSFLDPGSGGYSWQDAHADYGHVGLTGRVTEC 406
361 LSLDKLEAAPSEALVSRALQQLAEHFWAGVFLGPDSDPTHEPTDLPDGPCHVRIKI 420
407 LSLDKLEAAPSEALVSRALQQLAEHFWAGVFLGPDSDPTHEPTDLPDGPCHVRIKI 466
421 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 480
467 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 526
481 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 540

Db 527 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 586
Qy 541 SRAVLWGLWFLSCIGPFLLSAALLVVLKLDLIPYSHPGVFLFLAAFAVAVTQSFL 600
Db 587 SRAVLWGLWFLSCIGPFLLSAALLVVLKLDLIPYSHPGVFLFLAAFAVAVTQSFL 646
Qy 601 SAFPSRANLAAACGGLAYFSLYLYVLCVAVRDRLPAGGRVAASLLSPVAFGFGCSLAL 660
Db 647 SAFPSRANLAAACGGLAYFSLYLYVLCVAVRDRLPAGGRVAASLLSPVAFGFGCSLAL 706
Qy 661 LEBQSGEAGWNNVGTPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPMN 720
Db 707 LEBQSGEAGWNNVGTPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPMN 766
Qy 721 FPPRRSYWCGPRPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 767 FPPRRSYWCGPRPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRG 826
Qy 781 LSLDFYQGHITATFLGHNGAGKTTTSLTSLGFLPPSGSAFLLGHVDRSSMAAIRPHLVC 840
Db 827 LSLDFYQGHITATFLGHNGAGKTTTSLTSLGFLPPSGSAFLLGHVDRSSMAAIRPHLVC 886
Qy 841 PQYNVLFDMLTVDHEHWFYGRLLKGLSAAVVGPDQRLLDQVGLVSKSVQVTRHLSGMQR 900
Db 887 PQYNVLFDMLTVDHEHWFYGRLLKGLSAAVVGPDQRLLDQVGLVSKSVQVTRHLSGMQR 946
Qy 901 KLSVAIAFVGGSVWILDEPTAGVDPASRRGIWELLKXREGRTLLSTHHLDEAELLGD 960
Db 947 KLSVAIAFVGGSVWILDEPTAGVDPASRRGIWELLKXREGRTLLSTHHLDEAELLGD 1006
Qy 961 RVAVVAGGRLLCCCGSPLEFRRHLSGGVYLTVKARLPLTTNEKADTMEGSDVTROBKN 1020
Db 1007 RVAVVAGGRLLCCCGSPLEFRRHLSGGVYLTVKARLPLTTNEKADTMEGSDVTROBKN 1066
Qy 1021 GSGSRVGTPTQALLALVQHWVPGARLVEELPHELVLPYTGARHGDGFSATLIFRELDTRAE 1080
Db 1067 GSGSRVGTPTQALLALVQHWVPGARLVEELPHELVLPYTGARHGDGFSATLIFRELDTRAE 1126
Qy 1081 LRLTGYSISTSLSEEIFLKVVEECAADTMEGSCGHLCTGIAGLDVTLRLKMPPOETA 1140
Db 1127 LRLTGYSISTSLSEEIFLKVVEECAADTMEGSCGHLCTGIAGLDVTLRLKMPPOETA 1186
Qy 1141 LENGEPAAGSAPETDQSGPDVAGRVGQWALTROQLQALLKRFLARRSRGFLFAQIVLP 1200
Db 1187 LENGEPAAGSAPETDQSGPDVAGRVGQWALTROQLQALLKRFLARRSRGFLFAQIVLP 1246
Qy 1201 ALFVGLALVPSLIVPPFGHYPALRLSPTMYGAQVSPFSEDAPGDPGRARLLEALLQEAGL 1260
Db 1247 ALFVGLALVPSLIVPPFGHYPALRLSPTMYGAQVSPFSEDAPGDPGRARLLEALLQEAGL 1306
Qy 1261 EEPVQVHSHRFGAPAEVPAEVAKVILASGNWTPESPACQSQPGARRLLPDCPAAAGGP 1320
Db 1307 EEPVQVHSHRFGAPAEVPAEVAKVILASGNWTPESPACQSQPGARRLLPDCPAAAGGP 1366
Qy 1321 PPQAVTGSVEVQNLTGRNLSDFLVKTYPRVQGLTKKKWNEVRYGFSLGRDPGL 1380
Db 1367 PPQAVTGSVEVQNLTGRNLSDFLVKTYPRVQGLTKKKWNEVRYGFSLGRDPGL 1426
Qy 1381 PSQGLGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLLDAQDSLKIFWNNKGNHSMVAF 1440
Db 1427 PSQGLGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLLDAQDSLKIFWNNKGNHSMVAF 1486
Qy 1441 VNRASNAILRAHLPDGPGRHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM 1500
Db 1487 VNRASNAILRAHLPDGPGRHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM 1546
Qy 1501 SFVPASFTLVLIERVTRAKHLQMGGLSPTLYWLGNFMDMCMNLYVPACIVLILFLAQ 1560
Db 1547 SFVPASFTLVLIERVTRAKHLQMGGLSPTLYWLGNFMDMCMNLYVPACIVLILFLAQ 1606
Qy 1561 QRAYVAPANLPAIILLLLLYGSITLPMYPASFFSVPTAYVVLTCINLFIGINGSNAT 1620

Db 1607 QRAYVAPANLPAALLLLLLYGSWSTIPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT 1666
QY 1621 FVLELFSQKLOEVSRIKQVFLIIPFHCLGRGLLDVNRQAMADAFERLDGROFQSPUR 1680
Db 1667 FVLELFSQKLOEVSRIKQVFLIIPFHCLGRGLLDVNRQAMADAFERLDGROFQSPUR 1726
QY 1681 WEVVGKLLAMVIOGFLPFLTLLOHRSQQLPQPRVSLPLGGEDEDAVARERVVQG 1740
Db 1727 WEVVGKLLAMVIOGFLPFLTLLOHRSQQLPQPRVSLPLGGEDEDAVARERVVQG 1786
QY 1741 ATQGDVLRNLTKVYRGORPANDRLCLGTPPGECFGLGVNGAGKTSFRWVTGDTLA 1800
Db 1787 ATQGDVLRNLTKVYRGORPANDRLCLGTPPGECFGLGVNGAGKTSFRWVTGDTLA 1846
QY 1801 SRGEAVLAGHSVARPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
Db 1847 SRGEAVLAGHSVARPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVQTA 1906
QY 1861 GSGLARLGLSWYADRPAGTYSGNKRKLATALLVGDPAVVFLEDDPTTGMPSARRPLWN 1920
Db 1907 GSGLARLGLSWYADRPAGTYSGNKRKLATALLVGDPAVVFLEDDPTTGMPSARRPLWN 1966
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1967 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQPAFAVAAEFPGSELREAHGGRRLRFPOLPGRCALARVFGELAVHGAHGVEDF 2040
Db 2027 PAARSQPAFAVAAEFPGSELREAHGGRRLRFPOLPGRCALARVFGELAVHGAHGVEDF 2086
QY 2041 SVSQTMLBEVFLYFSKQKQEDTSEEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2100
Db 2087 SVSQTMLBEVFLYFSKQKQEDTSEEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2146

RESULT 2
AAU04483
ID AAU04483 standard; protein; 2146 AA.
XX
AC AAU04483;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human PD-ATP-binding cassette (PD-ABC) protein form #1.
XX
KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.
XX
OS Homo sapiens.
XX
FN WO200153490-A1.
XX
PD 26-JUL-2001.
XX
PF 23-JAN-2001; 2001WO-US002191.
XX
PR 24-JAN-2000; 2000US-0177889P.
PR 30-JUN-2000; 2000US-0215405P.
XX
PA (WARN) WARNER LAMBERT CO.
XX
FI Johns MA, Tafuri SR, Wang M;
XX
DR WPI; 2001-442259/47.
XX
DR N-P8DB; AAS08706.
XX
PT New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.

XX
PS
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

Claim 10; Page 54-62; 77pp; English.

The sequence represents human PD-ATP-binding cassette (PD-ABC) protein form 1. PD-ABC maps to chromosome 19p13.3 and is expressed in various tissues including spleen, thymus, peripheral blood leukocytes, bone marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to diagnose and treat cardiovascular disorders, inflammatory disorders, dyslipidaemia, epilepsy, diseases related to abnormal calcium flux, coronary artery disease, Tangier's disease, familial high-density lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease, insulin resistance, obesity, alcoholism, retinal degeneration, hypertension and vascular disease. The sequences are also used in drug screening assays

Sequence 2146 AA;

Query Match 99.9%; Score 10883; DB 4; Length 2146;

Beat Local Similarity 99.9%; Pred. No. 0;

Matches 2098; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKPLPSAGTVPMLQGLICVNNVNTCFPOLTPGEPGRLSNENDSLVSRL 60
Db 47 PPLEHHECHFPNKPLPSAGTVPMLQGLICVNNVNTCFPOLTPGEPGRLSNENDSLVSRL 106
QY 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPPLPPMLDVAELTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPPLPPMLDVAELTSLRT 166
QY 121 ESLGALGQAQEPHLSLEAAEDLAQELLARLSVELRALLORPGTSGPLELSEALCS 180
Db 167 ESLGALGQAQEPHLSLEAAEDLAQELLARLSVELRALLORPGTSGPLELSEALCS 226
QY 181 VRGSSVTGPSLWYEASDLMELVQBPESALPOSSSPACSELIGALDHPSLRLWR 240
Db 227 VRGSSVTGPSLWYEASDLMELVQBPESALPOSSSPACSELIGALDHPSLRLWR 286
QY 241 LKPLILGKLLFAPDTPFTRKLMQVNTFFELTLRLDVREVMELGPRIFTFMNDSSNVA 300
Db 287 LKPLILGKLLFAPDTPFTRKLMQVNTFFELTLRLDVREVMELGPRIFTFMNDSSNVA 346
QY 301 MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
Db 347 MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406
QY 361 LSLDKLEAAPSEALVSRALQLLAHEHFWAGVFLGPEDESDPTHEHPTDLPGRHVRIKI 420
Db 407 LSLDKLEAAPSEALVSRALQLLAHEHFWAGVFLGPEDESDPTHEHPTDLPGRHVRIKI 466
QY 421 RMDIDVVRTNKRDRFMDPGPAADPLTLDRYVWGGFYVLQDLVERAAVRVLSGANPRAG 480
Db 467 RMDIDVVRTNKRDRFMDPGPAADPLTLDRYVWGGFYVLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCVDDVFLRVLSRSLPLFLTLAMIYSVTLTVAVREKTRLDRTVRANGL 540
Db 527 LYLOQMPYPCVDDVFLRVLSRSLPLFLTLAMIYSVTLTVAVREKTRLDRTVRANGL 586
QY 541 SRVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPYSHPGVVFLLAFAFATVVTQSFL 600
Db 587 SRVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPYSHPGVVFLLAFAFATVVTQSFL 646
QY 601 SAFFSRANLAAACGLAYFSLYPLVLCVWRDLRAGGRVAAASLLSPVAFGFCESLAL 660
Db 647 SAFFSRANLAAACGLAYFSLYPLVLCVWRDLRAGGRVAAASLLSPVAFGFCESLAL 706
QY 661 LEEQGEQAQHNVTGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 720
Db 707 LEEQGEQAQHNVTGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 766
QY 721 PPFRRSYWCGRPKPSAPCPTPLDPKVLEAPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 767 PPFRRSYWCGRPKPSAPCPTPLDPKVLEAPGLSPGVSVRSLEKRPFGSPQPALRG 826

QY 781 LSLDFYOGHITAFI LHNAGAGKTTTSLISGLFPSPGSGAFILGHDRVSSMAAIRPHLGVC 840
DB 827 LSLDFYOGHITAFI LHNAGAGKTTTSLISGLFPSPGSGAFILGHDRVSSMAAIRPHLGVC 886
QY 841 PQYNVLFDMITVDEHVFYGRKGLSAAVVGPEQDRLQLDGLVSKQSVOTRHLSSGGWQR 900
DB 887 PQYNVLFDMITVDEHVFYGRKGLSAAVVGPEQDRLQLDGLVSKQSVQTRHLSSGGWQR 946
QY 901 KLSVAIAFVGSGQVILDEPTAGVDPPASRRGIWELLKLYREGRTLILSTHHLDAEELGD 960
DB 947 KLSVAIAFVGSGQVILDEPTAGVDPPASRRGIWELLKLYREGRTLILSTHHLDAEELGD 1006
QY 961 RVAVVAGGRILCCCGSPFLRRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTRQEKKN 1020
DB 1007 RVAVVAGGRILCCCGSPFLRRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTRQEKKN 1066
QY 1021 GSQSRVGTTPOLLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATLFRELDTRLAE 1080
DB 1067 GSQSRVGTTPOLLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATLFRELDTRLAE 1126
QY 1081 LRLTGYGISTSLBEEIFLKVVEECAADTMDGSCGQHLCTGTIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGYGISTSLBEEIFLKVVEECAADTMDGSCGQHLCTGTIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAAGSAPETDQSGPDAGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
DB 1187 LENGEPAAGSAPETDQSGPDAGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1246
QY 1201 ALFVGLALVLSLIVPPPGHYPALRLSPMTYGAQVSFFSEADPGPRARLLEALLQAGL 1260
DB 1247 ALFVGLALVLSLIVPPPGHYPALRLSPMTYGAQVSFFSEADPGPRARLLEALLQAGL 1306
QY 1261 EEPVQVHSHRFSAPAEVPAEVAKVYLVASGNMTPESSPACQSQPGARRLLPDCPAAAGGP 1320
DB 1307 EEPVQVHSHRFSAPAEVPAEVAKVYLVASGNMTPESSPACQSQPGARRLLPDCPAAAGGP 1366
QY 1321 PPOAVTGSCEVQNLGTRNLSDFLVKTYPRLVQGLTKTKWNEVRYGGFSLGGRDPGL 1380
DB 1367 PPOAVTGSCEVQNLGTRNLSDFLVKTYPRLVQGLTKTKWNEVRYGGFSLGGRDPGL 1426
QY 1381 PSQGLGRSVEELWALLSPFGGALDRVLKNTAWAHSLSDAQDSLKIWFNNKGWHSVAF 1440
DB 1427 PSQGLGRSVEELWALLSPFGGALDRVLKNTAWAHSLSDAQDSLKIWFNNKGWHSVAF 1486
QY 1441 VNRSANILRAHLPPGPARHAHSITTLNHPUNLTKEOLESAALMASSVDVLVSIICVVFAM 1500
DB 1487 VNRSANILRAHLPPGPARHAHSITTLNHPUNLTKEOLESAALMASSVDVLVSIICVVFAM 1546
QY 1501 SFVPASFTLVLIIEERVTRAKHLQLMGGLSPTLYWGNFLDMCNLYLPACIVVLIFLAFQ 1560
DB 1547 SFVPASFTLVLIIEERVTRAKHLQLMGGLSPTLYWGNFLDMCNLYLPACIVVLIFLAFQ 1606
QY 1561 QRAYVAPANLALLLLLYGWSITPLMYPASFFSPVSTAYVVLTCINLFIGINGSMAT 1620
DB 1607 QRAYVAPANLALLLLLYGWSITPLMYPASFFSPVSTAYVVLTCINLFIGINGSMAT 1666
QY 1621 FVLEFSDQKLQEVSRILKQVFLIFPHFCIGRGLIDMVRNQAMADAFERLGDQFQSPFLR 1680
DB 1667 FVLEFSDQKLQEVSRILKQVFLIFPHFCIGRGLIDMVRNQAMADAFERLGDQFQSPFLR 1726
QY 1681 WEVVGKNLLAMVIOGFLPLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVAREERVVQG 1740
DB 1727 WEVVGKNLLAMVIOGFLPLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVAREERVVQG 1786
QY 1741 ATQGDVLVLRNLTKVYRGQMPAVDRICLGIIPGCECFGLLVNGAGKTSIFRWVGTDLA 1800
DB 1787 ATQGDVLVLRNLTKVYRGQMPAVDRICLGIIPGCECFGLLVNGAGKTSIFRWVGTDLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLGVPEAQAQTA 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLGVPEAQAQTA 1906
QY 1861 GSGLARLGLSWYADRPAAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLWN 1920

DB 1907 GSGLARLGLSWYADRPAAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLWN 1966
QY 1921 SLLAVVREGSRVMLTSHSMEECEALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLRLV 1980
DB 1967 SLLAVVREGSRVMLTSHSMEECEALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLRLV 2026
QY 1981 PAARSQPAALFVAABEPPGSELREAHGRLRFQLPQGRCLARVFGELAVHGAHGVEDF 2040
DB 2027 PAARSQPAALFVAABEPPGSELREAHGRLRFQLPQGRCLARVFGELAVHGAHGVEDF 2086
QY 2041 SVSQTMLVEEVFLYPSKQDQKDETEQKEAGVGVDPAQGLQHPKRVSQFLDDSTAEITVL 2100
DB 2087 SVSQTMLVEEVFLYPSKQDQKDETEQKEAGVGVDPAQGLQHPKRVSQFLDDSTAEITVL 2146
RESULT 3
AAO14210
ID AAO14210 standard; protein; 2180 AA.
XX AAO14210;
AC AAO14210;
XX
DT 03-MAY-2002 (first entry)
XX
XX Human transporter and ion channel TRICH-27.
DE
XX
XX Human; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder;
KW cell proliferative disorder; neuroprotective; nootropic;
KW cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX WO200204520-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US021448.
XX
XX 07-JUL-2000; 2000US-0216547P.
PR 14-JUL-2000; 2000US-0218232P.
PR 21-JUL-2000; 2000US-0220112P.
PR 28-JUL-2000; 2000US-0221839P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J;
PI Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Wallia NK; Lu Y;
PI Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
PI Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;
PI Kearney L, Thangavelu K, Das D, Policky JL;
XX
XX WPI; 2002-205969/26.
DR N-PSDB; AAL44693.
XX
XX New human transporters and ion channel polypeptides for diagnosing,
PT treating or preventing transport, neurological, muscle, immunological and
PT cell proliferative disorders.
XX
XX Claim 1; Page 188-192; 230pp; English.
PS
XX The present invention provides the protein and coding sequences of a
CC number of human transporter and ion channel proteins, designated TRICH-1-
CC TRICH-32. The sequences can be used in the treatment of transport,
CC neurological, muscle, immunological and cell proliferative disorders. The
CC present sequence is a protein of the invention
XX
XX Sequence 2180 AA;
SQ
Query Match 99.9%; Score 10883; DB 5; Length 2180;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	PPLEHCHFFNKPPLPSAGTVPWLOGLICNVNNTCFPOLTPGEEPCGRLSNFNDSLSVRL	60
Db	81	PPLEHCHFFNKPPLPSAGTVPWLOGLICNVNNTCFPOLTPGEEPCGRLSNFNDSLSVRL	140
Qy	61	ADARTVLGASAHRTLAGLKGKIATLRAARSTAOQPTKQSPLEPMPMDVABELLTSLRT	120
Db	141	ADARTVLGASAHRTLAGLKGKIATLRAARSTAOQPTKQSPLEPMPMDVABELLTSLRT	200
Qy	121	ESLGALGQAQPLHSLLEAEDLAQELLALSLVELRALLORPGTSGPLELSEALCS	180
Db	201	ESLGALGQAQPLHSLLEAEDLAQELLALSLVELRALLORPGTSGPLELSEALCS	260
Qy	181	VRGPGSTVGPSSLNWTAEADLMELVQEPESALPDSLSPACSELIGALDHPLSLLWRR	240
Db	261	VRGPGSTVGPSSLNWTAEADLMELVQEPESALPDSLSPACSELIGALDHPLSLLWRR	320
Qy	241	LKPLILGKLLFAPDTPFTRKMAQVNRTPFEELTLRDVREVMELGPRIFTFMDSSNVA	300
Db	321	LKPLILGKLLFAPDTPFTRKMAQVNRTPFEELTLRDVREVMELGPRIFTFMDSSNVA	380
Qy	301	MLORLQWODEGRRPRGGRDHMEALRSFLDPGSGGYSWODAHADVGHVCTIGRVTEC	360
Db	381	MLORLQWODEGRRPRGGRDHMEALRSFLDPGSGGYSWODAHADVGHVCTIGRVTEC	440
Qy	361	LSLDKLEAPSAALVSRAQLLAEHRFWAGVVFILGPEDSDPTBHTPDLPQGHVRIKI	420
Db	441	LSLDKLEAPSAALVSRAQLLAEHRFWAGVVFILGPEDSDPTBHTPDLPQGHVRIKI	500
Qy	421	RMDIDVTRTKIRDRFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	480
Db	501	RMDIDVTRTKIRDRFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	560
Qy	481	LYLOQMPYCYDDVFLVLSLSLPLFLTLAMIYSVTLTKAVVREKETRLDTRAMGL	540
Db	561	LYLOQMPYCYDDVFLVLSLSLPLFLTLAMIYSVTLTKAVVREKETRLDTRAMGL	620
Qy	541	SRVILWLGWFLSCIGPFLLSALLVLVLKLGIDILPYSHPGVVFLLAFAFNAVTVQSFL	600
Db	621	SRVILWLGWFLSCIGPFLLSALLVLVLKLGIDILPYSHPGVVFLLAFAFNAVTVQSFL	680
Qy	601	SAFFSRANLAAACGLAYFSLYLPVYLCVAMWRDLRPAAGRVAASLLSPVAFGFCESLAL	660
Db	681	SAFFSRANLAAACGLAYFSLYLPVYLCVAMWRDLRPAAGRVAASLLSPVAFGFCESLAL	740
Qy	661	LEEQEGQAQWNVGTRPTADVFSLAQVSGLLLLDAAALGLATWYLEAVCPQGYGIPBPWN	720
Db	741	LEEQEGQAQWNVGTRPTADVFSLAQVSGLLLLDAAALGLATWYLEAVCPQGYGIPBPWN	800
Qy	721	PPFRSRYWCGPRPPKSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPQPALRG	780
Db	801	PPFRSRYWCGPRPPKSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPQPALRG	860
Qy	781	LSLDFYQGHITAFHGNGAGKTTLSLSGLPPSPGSAFLLGHVDVRSMAAIRPHLGYC	840
Db	861	LSLDFYQGHITAFHGNGAGKTTLSLSGLPPSPGSAFLLGHVDVRSMAAIRPHLGYC	920
Qy	841	POYNVLFDMLTVDHEHVFYGRUKGLSAAVVGPEQDRLLQDVLGSKQSVQTRHLSGQMOR	900
Db	921	POYNVLFDMLTVDHEHVFYGRUKGLSAAVVGPEQDRLLQDVLGSKQSVQTRHLSGQMOR	980
Qy	901	KLVSIAIAFVGGSQVWILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEALLGD	960
Db	981	KLVSIAIAFVGGSQVWILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEALLGD	1040
Qy	961	RVAVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKKN	1020
Db	1041	RVAVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKKN	1100
Qy	1021	GSQGSRVGTPQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFRDLTRLAE	1080
Db	1101	GSQGSRVGTPQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFRDLTRLAE	1160

RESULT 4

ABP52096

ID ABP52096 standard; protein; 2146 AA.

Qy	1081	LRLTGYSIDTSLEBEIFLKVVVEECAADTDMEDGSCQHLCTGTGIAGLDVTLRLKMPQETA	1140
Db	1161	LRLTGYSIDTSLEBEIFLKVVVEECAADTDMEDGSCQHLCTGTGIAGLDVTLRLKMPQETA	1220
Qy	1141	LENBEPAGSABETDGGSDPDVGRVQGWALTRQLOALLKRLFLARSRRLGFAQIVLP	1200
Db	1221	LENBEPAGSABETDGGSDPDVGRVQGWALTRQLOALLKRLFLARSRRLGFAQIVLP	1280
Qy	1201	ALFVGLALVFSLIYVPPFGHYPALRLSPMYGAQVSFFSEADPGDPGRARLLLEALLQEAGL	1260
Db	1281	ALFVGLALVFSLIYVPPFGHYPALRLSPMYGAQVSFFSEADPGDPGRARLLLEALLQEAGL	1340
Qy	1261	BEPPVQHSHRFSAPEVPAEYAKVLASGNWTPESPSPACQCSQPCQGRARLLPDCPAAAGP	1320
Db	1341	BEPPVQHSHRFSAPEVPAEYAKVLASGNWTPESPSPACQCSQPCQGRARLLPDCPAAAGP	1400
Qy	1321	PPQAVTSGGEVQVNLTORNLSDFLVTKTYPRLVROGLTKKWNNEVRVGGFSLGDRDGL	1380
Db	1401	PPQAVTSGGEVQVNLTORNLSDFLVTKTYPRLVROGLTKKWNNEVRVGGFSLGDRDGL	1460
Qy	1381	PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLEDAQDSLKIWFNNKGWSMAF	1440
Db	1461	PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLEDAQDSLKIWFNNKGWSMAF	1520
Qy	1441	VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKQELSEALMASSVDVLVSI CVVFAM	1500
Db	1521	VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKQELSEALMASSVDVLVSI CVVFAM	1580
Qy	1501	SFVPASFTVLVIERVTRAKHLQLMGGLSPTLYMLGNFLMDMNCNLYLPACIVLILFAPQ	1560
Db	1581	SFVPASFTVLVIERVTRAKHLQLMGGLSPTLYMLGNFLMDMNCNLYLPACIVLILFAPQ	1640
Qy	1561	QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAVVLTCLNLFITGINGSMAT	1620
Db	1641	QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAVVLTCLNLFITGINGSMAT	1700
Qy	1621	FVLELFSQKLEQSVSRILKQVFLIPPHFCLAGRLIDMVRNQAMADAFERLGRDQFQSPUR	1680
Db	1701	FVLELFSQKLEQSVSRILKQVFLIPPHFCLAGRLIDMVRNQAMADAFERLGRDQFQSPUR	1760
Qy	1681	WEVVGKILLAMVIOGFLPLLFTLLLOHRSQLLPOPRVRSLLPILGSEDEDAVARERVVQG	1740
Db	1761	WEVVGKILLAMVIOGFLPLLFTLLLOHRSQLLPOPRVRSLLPILGSEDEDAVARERVVQG	1820
Qy	1741	ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGCEFCGLGVNGAGKTTSTFRMVTGDTL	1800
Db	1821	ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGCEFCGLGVNGAGKTTSTFRMVTGDTL	1880
Qy	1801	SRGEAVLAGHSVAREPSSAHLISMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVOTA	1860
Db	1881	SRGEAVLAGHSVAREPSSAHLISMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVOTA	1940
Qy	1861	GSGLARLGLSNYADRPAGTYSGNKRKLATLALVGDPAVFLDPTTGMDSABRRFLWN	1920
Db	1941	GSGLARLGLSNYADRPAGTYSGNKRKLATLALVGDPAVFLDPTTGMDSABRRFLWN	2000
Qy	1921	SLLAVVREGSRVMTLSHSMEECEALCSFLATMNGRFFCLGSPQHLKGRFAAGHTLTURV	1980
Db	2001	SLLAVVREGSRVMTLSHSMEECEALCSFLATMNGRFFCLGSPQHLKGRFAAGHTLTURV	2060
Qy	1981	PAARSQAAAAFVAABFPGSSELRHAGGRRLRQLPPGGRCALARVFGELAVHGAEGHVEDF	2040
Db	2061	PAARSQAAAAFVAABFPGSSELRHAGGRRLRQLPPGGRCALARVFGELAVHGAEGHVEDF	2120
Qy	2041	SVSOTMLEEVLVYFSKQKQKDETEEQEAGVVDPAFGLQHPKRVVSFLDPPSTAEVTL	2100
Db	2121	SVSOTMLEEVLVYFSKQKQKDETEEQEAGVVDPAFGLQHPKRVVSFLDPPSTAEVTL	2180

QY 1381 PSQELGRSVEELWALLSPGALDRVLKNTAWAHSLSDAQDSLKIWNKNGHSMVAF 1440
DB 1427 PSQELGRSVEELWALLSPGALDRVLKNTAWAHSLSDAQDSLKIWNKNGHSMVAF 1486
QY 1441 VNRSNAIILRAHLPPGPARHAHSITTLNHPILNLTKEQLSEAAIMASSVDVLVSIQVVFAM 1500
DB 1487 VNRSNAIILRAHLPPGPARHAHSITTLNHPILNLTKEQLFEAAIMASSVDVLVSIQVVFAM 1546
QY 1501 SFVPASFTLVLEERVTRAKHLQMGSLPTLYWLGNFMDMNCVLPACIVVLIFLAPO 1560
DB 1547 SFVPASFTLVLEERVTRAKHLQMGSLPTLYWLGNFMDMNCVLPACIVVLIFLAPO 1606
QY 1561 QRAYVAPANLPAALLLLLYGHSITPLMYPASFFSVSTAYVLTICNLFINGSMAT 1620
DB 1607 QRAYVAPANLPAALLLLLYGHSITPLMYPASFFSVSTAYVLTICNLFINGSMAT 1666
QY 1621 FVLELFSQKLOEVRIRILKQVFLIIPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1680
DB 1667 FVLELFSQKLOEVRIRILKQVFLIIPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1726
QY 1681 WNVGKNLLAMVIOGFLFLLTLLQHRSQLPQPRVSLPLGGEEDVARERERVQOG 1740
DB 1727 WNVGKNLLAMVIOGFLFLLTLLQHRSQLPQPRVSLPLGGEEDVARERERVQOG 1786
QY 1741 ATQGDVLVRLNLTKVYRGQMPAVDRLCLIGIPGECFGLLVGNAGKTTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLVRLNLTKVYRGQMPAVDRLCLIGIPGECFGLLVGNAGKTTSTFRMVTGDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTA 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTA 1906
QY 1861 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVVFLEDPPTTGMPSARRFLWN 1920
DB 1907 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVVFLEDPPTTGMPSARRFLWN 1966
QY 1921 SLLAVVREGSVMLTSHSMECEALCSRIAIMVNGRFRCLGSPQHLKGRFAAGHTLTIRV 1980
DB 1967 SLLAVVREGSVMLTSHSMECEALCSRIAIMVNGRFRCLGSPQHLKGRFAAGHTLTIRV 2026
QY 1981 PAARSQPAAAFVAAPFGSELREAHGRLRFOLPPGRCALARVFGELAVHGAHGVDF 2040
DB 2027 PAARSQPAAAFVAAPFGSELREAHGRLRFOLPPGRCALARVFGELAVHGAHGVDF 2086
QY 2041 SVSQTMLSEVFLYFSKQKQEDTTEQKEAGVGVDPAQLOHPKXVSQFLDDPSTAEVTL 2100
DB 2087 SVSQTMLSEVFLYFSKQKQEDTTEQKEAGVGVDPAQLOHPKXVSQFLDDPSTAEVTL 2146

RESULT 5

ABU08466
ID ABU08466 standard; protein; 2146 AA.
XX AC ABU08466;
XX DT 18-JUN-2003 (first entry)
XX DE Human ABCA-SSN protein.
XX KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory; ABCA-SSN;
KW antiarteriosclerotic.
XX OS Homo sapiens.
XX PN WO2003010315-A1.
XX PD 06-FEB-2003.
XX PF 24-JUL-2002; 2002WO-JP007487.

XX 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX Claim 18; Page 164-174; 183pp; Japanese.
XX The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases. Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA-SSN protein
XX Sequence 2146 AA;
Query Match 99.9%; Score 10881; DB 6; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2098; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 PPLEHHECHFPNKPSPSAGTVPMQLGLICNVNNTCFPOLTGCEEPGRLSNFNDSLSVRL 60
DB 47 PPLEHHECHFPNKPSPSAGTVPMQLGLICNVNNTCFPOLTGCEEPGRLSNFNDSLSVRL 106
QY 61 ADARTVLGCGASAHRTLAGLGLIATLRAARSTAGPQPTKQSPLEPPMLDVAELLTSLRT 120
DB 107 ADARTVLGCGASAHRTLAGLGLIATLRAARSTAGPQPTKQSPLEPPMLDVAELLTSLRT 166
QY 121 ESILGALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGSGPLELSEALCS 180
DB 167 ESILGALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGSGPLELSEALCS 226
QY 181 VRGFSSTVGPLNMYEASDLMEVGOBPESALPOSSISPACSELIGALDSHPLSLNWR 240
DB 227 VRGFSSTVGPLNMYEASDLMEVGOBPESALPOSSISPACSELIGALDSHPLSLNWR 286
QY 241 LKPLILGKLLFAPDTPPTTRKLMQVNRFTFELTLRLDVRVEMVLMGPRIFTFMDSSNVA 300
DB 287 LKPLILGKLLFAPDTPPTTRKLMQVNRFTFELTLRLDVRVEMVLMGPRIFTFMDSSNVA 346
QY 301 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
DB 347 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406
QY 361 LSLDKLEAAPSEAAALVSRALQLLAHFRWAGVVFGLPEDSDSDPTEHPTDLPQGHVRIKI 420
DB 407 LSLDKLEAAPSEAAALVSRALQLLAHFRWAGVVFGLPEDSDSDPTEHPTDLPQGHVRIKI 466
QY 421 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCVDDVFLRVLRSPLFLTLAWIYSVTLTVKAVVREKTRLDTRWRMGL 540
DB 527 LYLOQMPYPCVDDVFLRVLRSPLFLTLAWIYSVTLTVKAVVREKTRLDTRWRMGL 586
QY 541 SRAVLMLGWFLSCIGPPLLSAALLVVLKLGDIILPYSHPGVVFLLFAFAFATVQTQSFLL 600
DB 587 SRAVLMLGWFLSCIGPPLLSAALLVVLKLGDIILPYSHPGVVFLLFAFAFATVQTQSFLL 646
QY 601 SAFFSRANLAACGGLAYFSLYLPVLCVWRDLRDPAGGRVAASLLSPVAFGFCESLAL 660

Db 647 SAPFSRANLAAACGLAYFSLYLVLCVAVWRDLRPAAGRVAAASLLSPVAFGFCESIAL 706
Qy 661 LEEQGEQAQHNVTGRTADVFSIAQVSGLLLLDAAALYGLATWYLEAVCPQYIPEPNW 720
Db 707 LEEQGEQAQHNVTGRTADVFSIAQVSGLLLLDAAALYGLATWYLEAVCPQYIPEPNW 766
Qy 721 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGSLPGVSVPSLEKRPFGSPQALRG 780
Db 767 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGSLPGVSVPSLEKRPFGSPQALRG 826
Qy 781 LSLDFYOGHITAFIHLNGAGKTTLSILSGLFPSPGGSFAFLGHDVRSMAAIRPHLGVC 840
Db 827 LSLDFYOGHITAFIHLNGAGKTTLSILSGLFPSPGGSFAFLGHDVRSMAAIRPHLGVC 886
Qy 841 PQYNVLFDMLTVDHVMFYGRKLSAAVVGPEODRLLDQVGLVSKOSVQTRHLSGQMOR 900
Db 887 PQYNVLFDMLTVDHVMFYGRKLSAAVVGPEODRLLDQVGLVSKOSVQTRHLSGQMOR 946
Qy 901 KLSVAIAFVCGSQVILDEPTAGVDPASRRGIWELLAKYREGRTLILSTHLDDEAELLGD 960
Db 947 KLSVAIAFVCGSQVILDEPTAGVDPASRRGIWELLAKYREGRTLILSTHLDDEAELLGD 1006
Qy 961 RVAVVAGRLCCCGSPFLRRHLGSGYYLTILVKARLPLTTNEKADTDMEGSVDTROEKKN 1020
Db 1007 RVAVVAGRLCCCGSPFLRRHLGSGYYLTILVKARLPLTTNEKADTDMEGSVDTROEKKN 1066
Qy 1021 GSQSRVGTQQLLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATILFRELDTRLAE 1080
Db 1067 GSQSRVGTQQLLALVQHVWPGARLVEELPHELVLVLPYTGADHGSFATILFRELDTRLAE 1126
Qy 1081 LRLTGYGISTLSLEEFLKVVEECAADTDMEDSCGHLCTGIAGLDVTLRLKWPQETA 1140
Db 1127 LRLTGYGISTLSLEEFLKVVEECAADTDMEDSCGHLCTGIAGLDVTLRLKWPQETA 1186
Qy 1141 LENEPAGSAPETDQSGGPDVAVGVQGWALTROQLQALLKRFLLARRSRRLGFAQIVLP 1200
Db 1187 LENEPAGSAPETDQSGGPDVAVGVQGWALTROQLQALLKRFLLARRSRRLGFAQIVLP 1246
Qy 1201 ALFVGLALVSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDPGRARLLEALLQAGL 1260
Db 1247 ALFVGLALVSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDPGRARLLEALLQAGL 1306
Qy 1261 EEPVVOHSSHRFSAPEVAEVAKLASGNMTPEPSPACCSQOPGARLLPDCPAAAGGP 1320
Db 1307 EEPVVOHSSHRFSAPEVAEVAKLASGNMTPEPSPACCSQOPGARLLPDCPAAAGGP 1366
Qy 1321 PPOAVTGSGEVQNLITGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGL 1380
Db 1367 PPOAVTGSGEVQNLITGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGL 1426
Qy 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSDAQSLKTIWFNNKGWHSWVAF 1440
Db 1427 PSGQELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSDAQSLKTIWFNNKGWHSWVAF 1486
Qy 1441 VNRASNAILRAHLPPGPARAHASITTLNHPNLNLTKEQLSEALMASSVDVLVSTCVVFAM 1500
Db 1487 VNRASNAILRAHLPPGPARAHASITTLNHPNLNLTKEQLFEAALMASSVDVLVSTCVVFAM 1546
Qy 1501 SFVPASTFLVIEERVTRAKHLQMGGLSPTLYWLGFLWDMCNLYLPACIVVLIFLAFQ 1560
Db 1547 SFVPASTFLVIEERVTRAKHLQMGGLSPTLYWLGFLWDMCNLYLPACIVVLIFLAFQ 1606
Qy 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFPFSPSTAYVVLITCINLFIGINGSMAT 1620
Db 1607 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFPFSPSTAYVVLITCINLFIGINGSMAT 1666
Qy 1621 FVLELFSQDKLOEVSRLIKQVFLIFPHFCILGRGLIDMVRQAMADAPERIGDRQFQSPLR 1680
Db 1667 FVLELFSQDKLOEVSRLIKQVFLIFPHFCILGRGLIDMVRQAMADAPERIGDRQFQSPLR 1726
Qy 1681 WEVVGKNNLAWIOGPIFLFTLLLOHRSOLLPOPRVRSIPLLGEEDVARRERVVQG 1740
Db 1727 WEVVGKNNLAWIOGPIFLFTLLLOHRSOLLPOPRVRSIPLLGEEDVARRERVVQG 1786

RESULT 6

AAU09174

ID AAU09174 standard; protein; 2144 AA.

XX AAU09174;

AC AAU09174;

XX DT

09-APR-2002 (first entry)

XX DT

Human transporter molecule, MTP-1.

DE DE

XX DE

XX DE

KW Human; transporter molecule; MTP-1; cytostatic; Nootropic; HIV;

KW Neuroprotective; Antiparkinsonian; Anticonvulsant; Antianaemic; protein;

KW Antidiabetic; Antiartherosclerotic; Anti-human immunodeficiency virus;

KW Antiarthritic; Immunosuppressive; Antitubercular; Tuberculostatic;

KW Antiulcer; Antimanic; Tranquilliser; Vasotropic; fertility disorder;

KW transporter-associated disorder; haematopoietic disorder; anxiety;

KW leukocytic disorder; lipid metabolism; abnormal vascularisation;

KW immunological disorder; inflammatory disease; neurological disorder;

KW obsessive-compulsive disorder; cardiac-related disorder;

KW hormonal disorder; reproductive disorder.

XX Homo sapiens.

OS Homo sapiens.

XX WO200187978-A2.

XX WO200187978-A2.

XX 22-NOV-2001.

XX 22-NOV-2001.

XX 14-MAY-2001; 2001WO-US015533.

XX 14-MAY-2001; 2001WO-US015533.

XX 12-MAY-2000; 2000US-0204211P.

XX 12-MAY-2000; 2000US-0204211P.

XX (MILL-) MILLENNIUM PHARM INC.

XX (MILL-) MILLENNIUM PHARM INC.

XX Gluckmann M, Curtis RAJ;

XX Gluckmann M, Curtis RAJ;

XX WPI; 2002-082985/11.

XX WPI; 2002-082985/11.

XX N-PSDB; AAS19207.

XX N-PSDB; AAS19207.

XX New membrane transport protein and polynucleotides, useful for diagnosing

XX and treating transport protein related disorders e.g. cancer, restenosis,

XX asthma and Alzheimer's disease and to identify modulators of therapeutic

XX use.

XX Claim 12; Fig 1; 141pp; English.

XX The invention relates to an isolated membrane transporter protein-1 (MTP-

XX 1) (I). (I) is useful for identifying a compound which modulates the

New membrane transport protein and polynucleotides, useful for diagnosing and treating transport protein related disorders e.g. cancer, restenosis, asthma and Alzheimer's disease and to identify modulators of therapeutic use.

Claim 12; Fig 1; 141pp; English.

The invention relates to an isolated membrane transporter protein-1 (MTP-1) (I). (I) is useful for identifying a compound which modulates the

CC activity of (I). The method comprises contacting (I) or cell expressing
 CC (I) with a test compound and determining whether (I) binds to the test
 CC compound or determining the effect of the compound on the activity or
 CC expression of (I). The identified compound is useful in treatment and
 CC diagnosis of a subject having disorders characterised by aberrant or
 CC unwanted MTP-1 protein or nucleic acid expression or activity, where
 CC transporter-associated disorders include haematopoietic disorders,
 CC leukocytic disorders, disorders related to lipid metabolism, disorders
 CC involving abnormal vascularisation, immunological disorders, inflammatory
 CC diseases, neurological disorders, anxiety disorders, obsessive-compulsive
 CC disorders, cardiac-related disorders. Disorders also include cellular
 CC proliferation, growth, differentiation, hormonal disorders and
 CC reproductive or fertility disorders. The present sequence represents the
 CC amino acid sequence of human transporter molecule, MTP-1
 XX Sequence 2144 AA;

Query Match 99.9%; Score 10880; DB 5; Length 2144;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKPSPAGTVPWLOGLICNVNNTCPQPTGCEEPGRLSNFNDSLSVRL 60
 DB 45 PPLEHHECHFPNKPSPAGTVPWLOGLICNVNNTCPQPTGCEEPGRLSNFNDSLSVRL 104
 QY 61 ADARTVLGASAHRTLAGLGLKLIATRAARSTAPQPTKQSPLEPPMLDVAELLTSLRT 120
 DB 105 ADARTVLGASAHRTLAGLGLKLIATRAARSTAPQPTKQSPLEPPMLDVAELLTSLRT 164
 QY 121 ESLGALGAQAPPLSLLEAARDLAEALLALRSVLRELALORPGTSGPIELSEALCS 180
 DB 165 ESLGALGAQAPPLSLLEAARDLAEALLALRSVLRELALORPGTSGPIELSEALCS 224
 QY 181 VRGSSSTVGPSLNWYEAADLMELVQEPESALPDSSLPACSELIGALDHPLELLWRR 240
 DB 225 VRGSSSTVGPSLNWYEAADLMELVQEPESALPDSSLPACSELIGALDHPLELLWRR 284
 QY 241 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEELTLRDVREYWEMLGPRIFTFMDNSSVA 300
 DB 285 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEELTLRDVREYWEMLGPRIFTFMDNSSVA 344
 QY 301 MLORLLQWQDEGRPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
 DB 345 MLORLLQWQDEGRPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 404
 QY 361 LSLDKLEAAPSEAAALVSALQLLAHRFWAGVFLGPEDSSDPTHEPTDILGPGHVRKI 420
 DB 405 LSLDKLEAAPSEAAALVSALQLLAHRFWAGVFLGPEDSSDPTHEPTDILGPGHVRKI 464
 QY 421 RMDIDVTRTNKIRDRFWDPGPAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
 DB 465 RMDIDVTRTNKIRDRFWDPGPAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 524
 QY 481 LYLOMPYPCVDDVFLVLSRSLPLFUTLAWIYSVTITKAVVREKETRLDRTRAMGL 540
 DB 525 LYLOMPYPCVDDVFLVLSRSLPLFUTLAWIYSVTITKAVVREKETRLDRTRAMGL 584
 QY 541 SRAVLWLGWFLSCLGPFLLSALLVLKLGDIILPYSHPGVVFLLAFAFATVTSPELL 600
 DB 585 SRAVLWLGWFLSCLGPFLLSALLVLKLGDIILPYSHPGVVFLLAFAFATVTSPELL 644
 QY 601 SAFFSRANLAAACGGLAYFSLYLPVLCVAVNRDLPPAGGRVAASLLSPVAFGFCESLAL 660
 DB 645 SAFFSRANLAAACGGLAYFSLYLPVLCVAVNRDLPPAGGRVAASLLSPVAFGFCESLAL 704
 QY 661 LEEQEGQWNVNTRPTADVPSLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPNW 720
 DB 705 LEEQEGQWNVNTRPTADVPSLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPNW 764
 QY 721 PFERRSYWCGRPKSPAPCPTPLDPKVLVEAPGLSPGVSVRSLEKRPFGSPQPALRG 780
 DB 765 PFERRSYWCGRPKSPAPCPTPLDPKVLVEAPGLSPGVSVRSLEKRPFGSPQPALRG 824

QY 781 LSLDFYQGHITAFILGHNGAGKTTTILSILSGLPPPGSGSAFILGHDRSSMAAIRPHLVC 840
 DB 825 LSLDFYQGHITAFILGHNGAGKTTTILSILSGLPPPGSGSAFILGHDRSSMAAIRPHLVC 884
 QY 841 POYNVLFDMLTVDHWHFYGRKJLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMOR 900
 DB 885 POYNVLFDMLTVDHWHFYGRKJLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMOR 944
 QY 901 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIIWELLKYREGRTILSTHLLDEALLGD 960
 DB 945 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIIWELLKYREGRTILSTHLLDEALLGD 1004
 QY 961 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1020
 DB 1005 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1064
 QY 1021 GSQSRVGTPLLALVQHWVPGARLVBELPHELVLVLPYTGADGSGFATLPLELDTRLAE 1080
 DB 1065 GSQSRVGTPLLALVQHWVPGARLVBELPHELVLVLPYTGADGSGFATLPLELDTRLAE 1124
 QY 1081 LRLTGYGSDTSLBEIFLKVVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1140
 DB 1125 LRLTGYGSDTSLBEIFLKVVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1184
 QY 1141 LENGEPAGSAPETDQSGGPDVGRVQGWALTRQQLQALLKFLFLARRSRRLGFAQIVLP 1200
 DB 1185 LENGEPAGSAPETDQSGGPDVGRVQGWALTRQQLQALLKFLFLARRSRRLGFAQIVLP 1244
 QY 1201 ALFVGLALVFSLIVPPFGHYPALRLSPITMYGAQVSFFSEDPDGPGRARLLBALLOEAGL 1260
 DB 1245 ALFVGLALVFSLIVPPFGHYPALRLSPITMYGAQVSFFSEDPDGPGRARLLBALLOEAGL 1304
 QY 1261 EEPVQVHSHRFSFAPEVPAEVAKVLASGNWTPESPACOCQOPCARLLPDCPAAAGP 1320
 DB 1305 EEPVQVHSHRFSFAPEVPAEVAKVLASGNWTPESPACOCQOPCARLLPDCPAAAGP 1364
 QY 1321 PPPQAVTSGSVEVQNLNLTGRNLSDFLVKTYTPRLVROGLTKKWNVNRVYGGSLGRDPGL 1380
 DB 1365 PPPQAVTSGSVEVQNLNLTGRNLSDFLVKTYTPRLVROGLTKKWNVNRVYGGSLGRDPGL 1424
 QY 1381 PSGQLGRSVBELWALLSLPLPGGALDRVLKNTAWAHSLEDAQDSLKIWNFNKGHSMVAF 1440
 DB 1425 PSGQLGRSVBELWALLSLPLPGGALDRVLKNTAWAHSLEDAQDSLKIWNFNKGHSMVAF 1484
 QY 1441 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEOLSEALMASSVDVLSICVVFAM 1500
 DB 1485 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEOLSEALMASSVDVLSICVVFAM 1544
 QY 1501 SFVPASFTVLIEBRVTRAKHLQMLGGLSPTLYWLNFLWDMCNVLPACIVVLI FLAFAQ 1560
 DB 1545 SFVPASFTVLIEBRVTRAKHLQMLGGLSPTLYWLNFLWDMCNVLPACIVVLI FLAFAQ 1604
 QY 1561 QRAYVAPANLPALELLLYGWSITPLMYPASFFSVPSTAYVVLTCINLFIGINGSMAT 1620
 DB 1605 QRAYVAPANLPALELLLYGWSITPLMYPASFFSVPSTAYVVLTCINLFIGINGSMAT 1664
 QY 1621 FVLELFSQDLQVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGRDQFQSPUR 1680
 DB 1665 FVLELFSQDLQVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGRDQFQSPUR 1724
 QY 1681 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQVRSLPLLGDEEDVARERERVVQG 1740
 DB 1725 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQVRSLPLLGDEEDVARERERVVQG 1784
 QY 1741 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGECFGLLVNGAGKTSFTFMVGTDTLA 1800
 DB 1785 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGECFGLLVNGAGKTSFTFMVGTDTLA 1844
 QY 1801 SRGAVALAGHSVAREPSAAHLSMGCYCQSDAI FELLTGREHLELLARLGRVPEQAQVATA 1860
 DB 1845 SRGAVALAGHSVAREPSAAHLSMGCYCQSDAI FELLTGREHLELLARLGRVPEQAQVATA 1904
 QY 1861 GSGLARLGLSWYADRAGTYSGGNKRKLATALALVGPVAVVFLDEPTTGMPSARRFLWN 1920

Db 1905 GSGLARGLSWADRPAGTSGGNKRLATLALVGDPAVFLDEPTTGNDPSARRELWN 1964
Qy 1921 SLIAVREGSVMLTSHMBECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1965 SLIAVREGSVMLTSHMBECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2024
Qy 1981 PAARSQAAAFAVAEPFGSELREAHGGRIRFOLPPGRCALARVFGELAVHGAHGVEDF 2040
Db 2025 PAARSQAAAFAVAEPFGSELREAHGGRIRFOLPPGRCALARVFGELAVHGAHGVEDF 2084
Qy 2041 SVSQTMLIEVFLPSKQDKDETEOEKAGVGVDPAPGLQHPKRVSOFLDDPSTAETVL 2100
Db 2085 SVSQTMLIEVFLPSKQDKDETEOEKAGVGVDPAPGLQHPKRVSOFLDDPSTAETVL 2144

RESULT 7
ADD37429
ID ADD37429 standard; protein; 2144 AA.
XX AC ADD37429;
XX DT 15-JAN-2004 (first entry)
XX DE Human transporter MTP-1.
XX KW Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
KW diabetes.
XX OS Homo sapiens.
XX FN US2003143675-A1.
XX PD 31-JUL-2003.
XX PF 22-MAY-2002; 2002US-00154419.

XX 12-MAY-2000; 2000US-0204211P.
PR 29-JUN-2000; 2000US-0215376P.
PR 31-JUL-2000; 2000US-0221769P.
PR 19-SEP-2000; 2000US-0233790P.
PR 25-SEP-2000; 2000US-0235107P.
PR 05-OCT-2000; 2000US-0238336P.
PR 14-NOV-2000; 2000US-0248364P.
PR 15-NOV-2000; 2000US-0248878P.
PR 18-DEC-2000; 2000US-0256588P.
PR 21-DEC-2000; 2000US-0258028P.
PR 22-JAN-2001; 2001US-0263169P.
PR 14-MAY-2001; 2001US-00858194.
PR 29-JUN-2001; 2001US-00895811.
PR 31-JUL-2001; 2001US-00919781.
PR 19-SEP-2001; 2001US-00957664.
PR 25-SEP-2001; 2001US-00964295.
PR 05-OCT-2001; 2001US-00972724.
PR 14-NOV-2001; 2001US-00002769.
PR 17-DEC-2001; 2001US-00024623.
PR 22-JAN-2002; 2002US-00055025.
XX (MILL-) MILLENNIUM PHARM INC.
XX Curtis RAJ, Glucksmann MA, Meyers RE;
PI WPI; 2003-851783/79.
XX N-PSDB; ADD37428, ADD37430.
XX New isolated nucleic acid, useful for preparing a composition for
FT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
PT or diabetes.
XX Claim 11; SEQ ID NO 2; 663pp; English.
PS
XX

CC The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.
XX
SQ Sequence 2144 AA;
Query Match 99.9%; Score 10880; DB 7; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPLEHECHFPNKLPSAGTVPWLQGLICNVNNTCPQLTPGPEGRGLSNFNDLSVRL 60
Db 45 PPLEHECHFPNKLPSAGTVPWLQGLICNVNNTCPQLTPGPEGRGLSNFNDLSVRL 104
Qy 61 ADARTVLGGASARTLAGIKLITATRAARSTAOPTKOSPLEPPMLDVAELLTSLRT 120
Db 105 ADARTVLGGASARTLAGIKLITATRAARSTAOPTKOSPLEPPMLDVAELLTSLRT 164
Qy 121 ESIGLALGQAQEPHLSLEAAEDLAQELLALRSLVELRALLQRPRTSGLELLSEALCS 180
Db 165 ESIGLALGQAQEPHLSLEAAEDLAQELLALRSLVELRALLQRPRTSGLELLSEALCS 224
Qy 181 VRGFSSTVGPSLNWYEASDLMELVQGPESALPDSLSLSPACSELIGALDHPHLSRLWRR 240
Db 225 VRGFSSTVGPSLNWYEASDLMELVQGPESALPDSLSLSPACSELIGALDHPHLSRLWRR 284
Qy 241 LKPLILGKLLFADPTPTFKLMAQVNTFEELTLRDVREVWEMGPRIETFWNDSNVA 300
Db 285 LKPLILGKLLFADPTPTFKLMAQVNTFEELTLRDVREVWEMGPRIETFWNDSNVA 344
Qy 301 MLQRLQMODGRRQRPGRDHMEALRSFLDPGSGYSWQDAHADVGHVLVGLGRVTEC 360
Db 345 MLQRLQMODGRRQRPGRDHMEALRSFLDPGSGYSWQDAHADVGHVLVGLGRVTEC 404
Qy 361 LSCLKLEAAPSEAAALVSRAQLLAEHRFWAGVVFGLPEDSSDTEHPTDPLGPHVRIKI 420
Db 405 LSCLKLEAAPSEAAALVSRAQLLAEHRFWAGVVFGLPEDSSDTEHPTDPLGPHVRIKI 464
Qy 421 RMDIDVVTRTNKIRDRFWDGPAADPLTDLYVWGGFVYLQDLVERAAVRLSGANPRAG 480
Db 465 RMDIDVVTRTNKIRDRFWDGPAADPLTDLYVWGGFVYLQDLVERAAVRLSGANPRAG 524
Qy 481 LYLQOMPYPYCVDDVFLRSLPLPLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 540
Db 525 LYLQOMPYPYCVDDVFLRSLPLPLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 584
Qy 541 SRAVLMLGWFLSLCLGPFLLSAALLVLVLKGLDILPYSHPGVWFLFAAFVAVTQSFLL 600
Db 585 SRAVLMLGWFLSLCLGPFLLSAALLVLVLKGLDILPYSHPGVWFLFAAFVAVTQSFLL 644
Qy 601 SAFFSRANLAAACGLLAYFSLYPLCYAVWRDLRPPAGGRVAAASLLSPVAFGFCESLAL 660
Db 645 SAFFSRANLAAACGLLAYFSLYPLCYAVWRDLRPPAGGRVAAASLLSPVAFGFCESLAL 704
Qy 661 LEEQGEAOMHNVTGRTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGPFWN 720
Db 705 LEEQGEAOMHNVTGRTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGPFWN 764
Qy 721 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 780
Db 765 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 824

QY 781 LSLDFYQGHITAFHGNGAGKTTTILSLSGLPPPGSGSAFILGHVDVRSSMAAIRPHLGV 840
DB 825 LSLDFYQGHITAFHGNGAGKTTTILSLSGLPPPGSGSAFILGHVDVRSSMAAIRPHLGV 884
QY 841 POYNVLFDMLTVDHWHVFGYRUKLGLSAAVVGPEQDRLLQDVLGSKQSVQTRHLGGMQR 900
DB 885 POYNVLFDMLTVDHWHVFGYRUKLGLSAAVVGPEQDRLLQDVLGSKQSVQTRHLGGMQR 944
QY 901 KLSVAIAFVGSGQVILDEPTAGVDPASRGRIWELLKYREGRTILSTHHLDEAELGCD 960
DB 945 KLSVAIAFVGSGQVILDEPTAGVDPASRGRIWELLKYREGRTILSTHHLDEAELGCD 1004
QY 961 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1020
DB 1005 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTROEKN 1064
QY 1021 GSQGRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRAE 1080
DB 1065 GSQGRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRAE 1124
QY 1081 LRLTGIGSDTLEIRIFLKVVEECAADTMEGSCGQHLCTGIGLDVTLRLKMPQETA 1140
DB 1125 LRLTGIGSDTLEIRIFLKVVEECAADTMEGSCGQHLCTGIGLDVTLRLKMPQETA 1184
QY 1141 LENGEPAAGPETDQSGPDVAVRGQWALTRQQLQALLKRLFLARRSRGLFAQIVLP 1200
DB 1185 LENGEPAAGPETDQSGPDVAVRGQWALTRQQLQALLKRLFLARRSRGLFAQIVLP 1244
QY 1201 ALFVGLALVFSILVPPFGHYPALRLSPMTYGAQVFFSDEADPGDPGRARLLLEALLQEAGL 1260
DB 1245 ALFVGLALVFSILVPPFGHYPALRLSPMTYGAQVFFSDEADPGDPGRARLLLEALLQEAGL 1304
QY 1261 BEPPVOHSHRSAPAEVAEVAKLASGNWTPESSPACQSQPCARLLPDCPAAAGP 1320
DB 1305 BEPPVOHSHRSAPAEVAEVAKLASGNWTPESSPACQSQPCARLLPDCPAAAGP 1364
QY 1321 PPPQAVTSGVEVQNLTKGNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSIGGRDPGL 1380
DB 1365 PPPQAVTSGVEVQNLTKGNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSIGGRDPGL 1424
QY 1381 PSGQLGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIWFNNKGHWSVAF 1440
DB 1425 PSGQLGRSVEELWALLSPLPGCALDRVLKNTAWAHSIDAQDSLKIWFNNKGHWSVAF 1484
QY 1441 VNRAASNATLRAHLPGRPARHAHSITTLNHPNLNTEQOLSEALMASSVDVLVSIQVVFAM 1500
DB 1485 VNRAASNATLRAHLPGRPARHAHSITTLNHPNLNTEQOLSEALMASSVDVLVSIQVVFAM 1544
QY 1501 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLGNFMDMNCNVLVPACIVVLIFLAFQ 1560
DB 1545 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLGNFMDMNCNVLVPACIVVLIFLAFQ 1604
QY 1561 QRAYVAPANLPAALLLLLYGWSITPLMYPASFFSVFSTAYVLTCLNLFIGINGSMAT 1620
DB 1605 QRAYVAPANLPAALLLLLYGWSITPLMYPASFFSVFSTAYVLTCLNLFIGINGSMAT 1664
QY 1621 FVLELPSQKQEVSRILKQVFLIPPHFCLGRGLIDMVRNQAMADAFERLGDQRQPSFLR 1680
DB 1665 FVLELPSQKQEVSRILKQVFLIPPHFCLGRGLIDMVRNQAMADAFERLGDQRQPSFLR 1724
QY 1681 WEVVGKLLAMVIOQFPLFTLLQHRSQLLPQPRVRSPLPGLGEDEDAVARERVVQG 1740
DB 1725 WEVVGKLLAMVIOQFPLFTLLQHRSQLLPQPRVRSPLPGLGEDEDAVARERVVQG 1784
QY 1741 ATQGDVLVRLNLTQYRQMPAVDRCLGIPPGECFGLLGVNGAGKTTSTFRMVTGDTLA 1800
DB 1785 ATQGDVLVRLNLTQYRQMPAVDRCLGIPPGECFGLLGVNGAGKTTSTFRMVTGDTLA 1844
QY 1801 SRGEAVLAGHSVAREPSPAHLISMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1860
DB 1845 SRGEAVLAGHSVAREPSPAHLISMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1904

RESULT 8

ABU54629
ID ABU54629 standard; protein; 2059 AA.
XX AC ABU54629;
XX DT 03-JUN-2003 (first entry)
XX Human NOVX polypeptide #88.
DE Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX OS Homo sapiens.
XX PN WO200281498-A2.
XX PD 17-OCT-2002.
XX 03-APR-2002; 2002WO-US010780.
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.

PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 21-NOV-2001; 2001US-0332779P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX (CURA-) CURAGEN CORP.
XX
XX Guo X. Kekuda R. Miller CE, Malyankar UM, Spytek KA;
PI Patrajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;
PI Gorman L, Shenoy SG, Pena CBA, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shamketa RA, Gangolli EA, Taupier RJ, Casman SU, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
XX WPI; 2003-046858/04.
DR N-PSDB; ABX72257.
XX
XX New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
XX Claim 1; Page 281-282; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
CC of the invention
XX
XX Sequence 2059 AA;
SQ
Query Match 95.4%; Score 10392.5; DB 6; Length 2059;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
QY 1 PPLEHHECHFPNKLPSAGTVPMLQGLICNVNNTCFPQLTGPEPGRLSNFNDLSVRL 60
DB 47 PPLEHHECHFPNKLPSAGTVPMLQGLICNVNNTCFPQLTGPEPGRLSNFNDLSVRL 106
QY 61 ADARTVLGGASAHRTLAGLGLKLIATLAARSTAOPTKQSPLEPPMLDVAELLTSLRT 120
DB 107 ADARTVLGGASAHRTLAGLGLKLIATLAARSTAOPTKQSPLEPPMLDVAELLTSLRT 166
QY 121 ESLGALGQAQEPHLSLEAAEDLAQELLALRSIVELRALLQRPGTSGPLELSEALCS 180
DB 167 ESLGALGQAQEPHLSLEAAEDLAQELLALRSIVELRALLQRPGTSGPLELSEALCS 226
QY 181 VRGSSSTVPSLNVEASDLMEVGQBPESALPDSSLSIPACSELIGALDHSPLSRLWRR 240
DB 227 VRGSSSTVPSLNVEASDLMEVGQBPESALPDSSLSIPACSELIGALDHSPLSRLWRR 286
QY 241 LKPLILGKLLFAPDTPTRKLMQVNTFEBLLTLLRDVREWEMGLPRIFTFMNDSSNVA 300
DB 287 LKPLILGKLLFAPDTPTRKLMQVNTFEBLLTLLRDVREWEMGLPRIFTFMNDSSNVA 346
QY 301 MLQRLLOQDDEGRQPRPGRDHMEALRSFLDPGSGGYSWQADAHVGLVGTIGRVTEC 360
DB 347 MLQRLLOQDDEGRQPRPGRDHMEALRSFLDPGSGGYSWQADAHVGLVGTIGRVTEC 406
QY 361 LSLDKLEAAPSEAAVSRALQLLAHRFWAGVFLGPEDSSDPTHEHTPDLGPGHVRKI 420

DB 407 LSLDKLEAAPSEAAVSRALQLLAHRFWAGVFLGPEDSSDPTHEHTPDLGPGHVRKI 466
QY 421 RMDIDVVTRTNKIRDRFMDPGPAADPLTDLRYVWGVFVYQDLVERAAVRLSGANPRAG 480
DB 467 RMDIDVVTRTNKIRDRFMDPGPAADPLTDLRYVWGVFVYQDLVERAAVRLSGANPRAG 526
QY 481 LYLOQMPYPCYVDDVFLRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRWGL 540
DB 527 LYLOQMPYPCYVDDVFLRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRWGL 586
QY 541 SRAVLWLGWFLSCIGPFLSAALLVLVLKGLDILPYSHPGVWFELAAFAVATVTSQFL 600
DB 587 SRAVLWLGWFLSCIGPFLSAALLVLVLKGLDILPYSHPGVWFELAAFAVATVTSQFL 646
QY 601 SAFFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLRPLAGGRVAAASLSPVAFGFCESIAL 660
DB 647 SAFFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLRPLAGGRVAAASLSPVAFGFCESIAL 706
QY 661 LEEQGEQAQWHNVGTRPTADVFSIAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 720
DB 707 LEEQGEQAQWHNVGTRPTADVFSIAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 766
QY 721 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
DB 767 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 826
QY 781 LSLDFYQGHITAFILGHNGAGKTTTSLILSGFPSPGSAFILGHVVRSSMAAIRPHLGV 840
DB 827 LSLDFYQGHITAFILGHNGAGKTTTSLILSGFPSPGSAFILGHVVRSSMAAIRPHLGV 886
QY 841 PQYNVLFDMLTVDHWFYGRILKLSAAVVPQDRLLQDVLGVSKOSVOTRHLSCGMQR 900
DB 887 PQYNVLFDMLTVDHWFYGRILKLSAAVVPQDRLLQDVLGVSKOSVOTRHLSCGMQR 946
QY 901 KLSVAITAFVCGSOWILDEPTAGVDPASRRGIWELLIKYREGRTLILSTHLEAEALLGD 960
DB 947 KLSVAITAFVCGSOWILDEPTAGVDPASRRGIWELLIKYREGRTLILSTHLEAEALLGD 1006
QY 961 RVAVVAGRLCCCGSPFLRRHLGSGYLLTVKARPLTTTNEKADTMEGSDVTRQSKKN 1020
DB 1007 RVAVVAGRLCCCGSPFLRRHLGSGYLLTVKARPLTTTNEKADTMEGSDVTRQSKKN 1066
QY 1021 GSQSRVGTPLLALVQHWVPGARLVEELPHELVLPYTGADHGSFATILFREDTFLAE 1080
DB 1067 GSQSRVGTPLLALVQHWVPGARLVEELPHELVLPYTGADHGSFATILFREDTFLAE 1126
QY 1081 LRLTGYISDTSLEEIPKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGYISDTSLEEIPKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAQSAPETDQSGGPDVAVGVQWALTRQOLQALLKRLFLARRRRLGPAQIVLP 1200
DB 1187 LENGEPAQSAPETDQSGGPDVAVGVQWALTRQOLQALLKRLFLARRRRLGPAQIVLP 1246
QY 1201 ALFVGLALVESLIVPPGHPYALRLSPTMYGAOVSEFEDAPGDPGRARLLLEALLQAGL 1260
DB 1247 ALFVGLALVESLIVPPGHPYALRLSPTMYGAOVSEFEDAPGDPGRARLLLEALLQAGL 1306
QY 1261 BEPPVQHSRHFSAPEVPAEVAKVASGNMTPSPSPACCSQFAGRLRLPDCPAAAGGP 1320
DB 1307 BEPPVQHSRHFSAPEVPAEVAKVASGNMTPSPSPACCSQFAGRLRLPDCPAAAGGP 1366
QY 1321 PPQAVTGSGEVQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSLGGRDPGL 1380
DB 1367 PPQAVTGSGEVQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSLGGRDPGL 1426
QY 1381 PSQOELGRSVEELWALLSLPGLGALDRVLKNLTAWAHSLLDAQSLTKWNNKGHSWVAF 1440
DB 1427 PSQOELGRSVEELWALLSLPGLGALDRVLKNLTAWAHSLLDAQSLTKWNNKGHSWVAF 1486
QY 1441 VNRASNAIRLAHLPFGPARHAHSITTNHPLNITKEQLSEAAALMASSVDLVSTCVVFAM 1500

Db 1487 VNRASNAIRALPPLPGPARHAHSITTLNHPNLTKQLSEAAALMASSVDVLVSICVVFAM 1546
QY 1501 SFVPASFVLVLEERVTRAKHLQLMGGLSPTLYWLGFLWDMCNVLPACIVVLIFLAFQ 1560
Db 1547 SFVPASFVLVLEERVTRAKHLQLMGGLSPTLYWLGFLWDM----- 1588
QY 1561 QRAYVAPANLPAALLLLLYGWSITPLMYPASFVSPSTAYVTLTCINLFIINGSMAT 1620
Db 1589 ----- 1588
QY 1621 FVLELFSDKLOEVSRIILKQVFLIPHFCLGRGLDMVRNOAWADAFERLGRQFQSPLR 1680
Db 1589 -----KLOEVSRIILKQVFLIPHFCLGRGLDMVRNOAWADAFERLGRQFQSPLR 1639
QY 1681 MEWVGKNLAMIYIQGFLPFLTLLOHRSQLLPQPRVRSPLPGLGEDEDAVARERERVQ 1740
Db 1640 MEWVGKNLAMIYIQGFLPFLTLLOHRSQLLPQPRVRSPLPGLGEDEDAVARERERVQ 1699
QY 1741 ATQGDVLVLRNLTKVYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSFRMVTGDTILA 1800
Db 1700 ATQGDVLVLRNLTKVYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSFRMVTGDTILA 1759
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
Db 1760 SRGEAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLELLARLGVPEAQVQTA 1819
QY 1861 GSGLARLGLSWADPAGTYSGNKRKLATALALVGDPAVFLDEPTTGMDSARRFLWN 1920
Db 1820 GSGLARLGLSWADPAGTYSGNKRKLATALALVGDPAVFLDEPTTGMDSARRFLWN 1879
QY 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1880 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1939
QY 1981 PAARQPAAFVAAFPFSGELBEAGGRRLRFOLPGRCALARVFGELAVHGAHGVEDF 2040
Db 1940 PAARQPAAFVAAFPFSGELBEAGGRRLRFOLPGRCALARVFGELAVHGAHGVEDF 1999
QY 2041 SVSQTMLEEVFLYFSKQDKEDTEQKEAGVGVDPAPGLQHPKRVSPFLDDPSTAEVTL 2100
Db 2000 SVSQTMLEEVFLYFSKQDKEDTEQKEAGVGVDPAPGLQHPKRVSPFLDDPSTAEVTL 2059

RESULT 9
ID ABU08464
XX ABU08464 standard; protein; 2008 AA.
AC ABU08464;
XX
DT 18-JUN-2003 (first entry)
XX
DE Amino acid sequence for human ABCA7 splice variant #1.
XX
KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT Region
FT 1..28
FT /note="Given as SEQ ID No:9 and specifically claimed in
FT Claim 2"
XX
XX WO2003010315-A1.
XX
XX 06-FEB-2003.
XX
XX 24-JUL-2002; 2002WO-JP007487.
XX
XX 25-JUL-2001; 2001JP-00224176.
XX

PR 06-DEC-2001; 2001JP-00372530.
XX (KYOM) KYOMA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
DR N-PSDB; ABX95283.
XX
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 1; Page 106-115; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #1
XX
SQ Sequence 2008 AA;
Query Match 94.2%; Score 10264; DB 6; Length 2008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 121 ESLGALGQAQEPHLSLEAAEDLAQELLALSLVELRALLQRPRTGSPLELLSEALCS 180
Db 29 ESLGALGQAQEPHLSLEAAEDLAQELLALSLVELRALLQRPRTGSPLELLSEALCS 88
QY 181 VRGSPSTVPSINWYEAADLMELVQBPESALPDSSLPACSELIIGALDHPSLRLMWR 240
Db 89 VRGSPSTVPSINWYEAADLMELVQBPESALPDSSLPACSELIIGALDHPSLRLMWR 148
QY 241 LKPLILGLKLLPAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIETFMNDSSNVA 300
Db 149 LKPLILGLKLLPAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIETFMNDSSNVA 208
QY 301 MLQRLLOMQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVLGTLGRVTEC 360
Db 209 MLQRLLOMQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVLGTLGRVTEC 268
QY 361 LSLDKLEAAPSEALVSRALQLLAEHRFWAGVVFGLGPDSSDPTPHPTDGLGPHVRIKI 420
Db 269 LSLDKLEAAPSEALVSRALQLLAEHRFWAGVVFGLGPDSSDPTPHPTDGLGPHVRIKI 328
QY 421 RMDIDVTRTKIRDRFWDGPAADPLTLDRVYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRDRFWDGPAADPLTLDRVYVWGGFVYLQDLVERAAVRVLSGANPRAG 388
QY 481 LYLOQMPYPCYVDDVFLRVLSRSLPFLTLAWIYSVTITVXAVVREKETRLDTRAMGL 540
Db 389 LYLOQMPYPCYVDDVFLRVLSRSLPFLTLAWIYSVTITVXAVVREKETRLDTRAMGL 448
QY 541 SRAVLWLGWFLSCIGPFLLSAALLVILKLDGILPYSHPGVVFLPLAFAVATVQSFLL 600
Db 449 SRAVLWLGWFLSCIGPFLLSAALLVILKLDGILPYSHPGVVFLPLAFAVATVQSFLL 508
QY 601 SAFPSRANLAACGLAYFSLYLPVLCVNRDLRIPAGGRVAASLSVPAFGGCSLAL 660
Db 509 SAFPSRANLAACGLAYFSLYLPVLCVNRDLRIPAGGRVAASLSVPAFGGCSLAL 568
QY 661 LEEQEGAQHNVGTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYGIPPEPN 720
Db 569 LEEQEGAQHNVGTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYGIPPEPN 628
QY 721 PFFRRSYWCGRRPKSPAPCTPDLDPKVLVEAPPGLSGVSVRSLEKRFPGSPQALRG 780
Db 629 PFFRRSYWCGRRPKSPAPCTPDLDPKVLVEAPPGLSGVSVRSLEKRFPGSPQALRG 688

QY 781 LSLDFYOGHITAFI LHNAGKATTTLSILSGLFPSPGSGSAFLLGHDPVSSMAAIRPHLGYC 840
Db 689 LSLDFYOGHITAFI LHNAGKATTTLSILSGLFPSPGSGSAFLLGHDPVSSMAAIRPHLGYC 748
QY 841 PQYNVLFDMITVDHVFYGRKGLGSAVVGPEQDRLLQDVLGYSKQSVOTRHLSSGMQR 900
Db 749 PQYNVLFDMITVDHVFYGRKGLGSAVVGPEQDRLLQDVLGYSKQSVOTRHLSSGMQR 808
QY 901 KLSVAIAFVGSQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDBAELLGD 960
Db 809 KLSVAIAFVGSQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDBAELLGD 868
QY 961 RVAVAGGRILCCCSPLFRHLGSGYLLTVKARLPLTNEKADTMEGSDVTRQSKN 1020
Db 869 RVAVAGGRILCCCSPLFRHLGSGYLLTVKARLPLTNEKADTMEGSDVTRQSKN 928
QY 1021 GSQSRVGTTPOLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLDTLAE 1080
Db 929 GSQSRVGTTPOLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLDTLAE 988
QY 1081 LRLTGYGISDTSLEEIFLKVVVECAADTMDGSCGQHLCTGTIAGLDVTLRLKMPQETA 1140
Db 989 LRLTGYGISDTSLEEIFLKVVVECAADTMDGSCGQHLCTGTIAGLDVTLRLKMPQETA 1048
QY 1141 LENGEPAAGSAPETDQSGPDVAGVQGNALTRQOLQALLIKRFLARRSRGLFAQIVLP 1200
Db 1049 LENGEPAAGSAPETDQSGPDVAGVQGNALTRQOLQALLIKRFLARRSRGLFAQIVLP 1108
QY 1201 ALFVGLALVFSLLIYVPPGCHYPALRLSPTMYGAQVSFFSEADPGPGRARLLEALLQAGL 1260
Db 1109 ALFVGLALVFSLLIYVPPGCHYPALRLSPTMYGAQVSFFSEADPGPGRARLLEALLQAGL 1168
QY 1261 EEPVQVSHSRFSAPEVPAEVAKVLASGNMTPEPSPACQSQPGARRLLPDCPAAAGGP 1320
Db 1169 EEPVQVSHSRFSAPEVPAEVAKVLASGNMTPEPSPACQSQPGARRLLPDCPAAAGGP 1228
QY 1321 PPOQAVTSGSEVQNLTRGNLSDFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGRDPGL 1380
Db 1229 PPOQAVTSGSEVQNLTRGNLSDFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGRDPGL 1288
QY 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSQAQDSLKTFWNNKGWHSVAF 1440
Db 1289 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSQAQDSLKTFWNNKGWHSVAF 1348
QY 1441 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAALMASSVDVLVSVVVFAM 1500
Db 1349 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAALMASSVDVLVSVVVFAM 1408
QY 1501 SFVPASFTVLIIERVTTRAKHLQMGGLSPTLYWLGNFMDMCMYLYPACIVVLIFLAQ 1560
Db 1409 SFVPASFTVLIIERVTTRAKHLQMGGLSPTLYWLGNFMDMCMYLYPACIVVLIFLAQ 1468
QY 1561 QRAVVAANLPAALLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMAT 1620
Db 1469 QRAVVAANLPAALLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMAT 1528
QY 1621 FVLELFSQDLQEVSRILKQVFLIPFHLFGRGLIDMVRNQAMADAFERLGDROFQSPLR 1680
Db 1529 FVLELFSQDLQEVSRILKQVFLIPFHLFGRGLIDMVRNQAMADAFERLGDROFQSPLR 1588
QY 1681 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQVRSLPLLEEEDVAREHRRVQ 1740
Db 1589 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQVRSLPLLEEEDVAREHRRVQ 1648
QY 1741 ATQGDVLVRLNLTKVYRGORMPADVRLCLGIPPECFCGLLGNGAGKTSIFRMVGTDLA 1800
Db 1649 ATQGDVLVRLNLTKVYRGORMPADVRLCLGIPPECFCGLLGNGAGKTSIFRMVGTDLA 1708
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCFQSDAIFELLTGREHLELLARLGRVPEAQVATA 1860
Db 1709 SRGEAVLAGHSVAREPSAAHLSMGYCFQSDAIFELLTGREHLELLARLGRVPEAQVATA 1768

QY 1861 GSGLARLGLSWYADRPAGTYSGNKRKLATALVGDPAVVFLDEPTTGMDSARRFLWN 1920
Db 1769 GSGLARLGLSWYADRPAGTYSGNKRKLATALVGDPAVVFLDEPTTGMDSARRFLWN 1828
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTV 1980
Db 1829 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTV 1888
QY 1981 PAARSQPAAFVAAEPFSGSELREAHGRLRFQLPPOGRCALARVFGELAVHGAHGVDF 2040
Db 1889 PAARSQPAAFVAAEPFSGSELREAHGRLRFQLPPOGRCALARVFGELAVHGAHGVDF 1948
QY 2041 SVSQTMLVEEFLYFSKQDQKDETEQKEAGVGDPAQGLQHPKRVSOQLDDPSTAEVL 2100
Db 1949 SVSQTMLVEEFLYFSKQDQKDETEQKEAGVGDPAQGLQHPKRVSOQLDDPSTAEVL 2008

RESULT 10
ABU08465
ID ABU08465 standard; protein; 1993 AA.
XX ABU08465;
AC AC
DT 18-JUN-2003 (first entry)
XX Amino acid sequence for human ABCA7 splice variant #2.
DE Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic.
XX Homo sapiens.
OS
XX
XX WO2003010315-A1.
PN
XX
PD 06-FEB-2003.
XX
PF 24-JUL-2002; 2002WO-JP007487.
XX
PR 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
XX Ueda K, Nakagawa S, Nagase T;
XX
DR WPI; 2003-239444/23.
DR N-PSDB; ABX95284.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 1; Page 145-154; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #2
XX
SQ Sequence 1993 AA;

Query Match 93.4%; Score 10171.5; DB 6; Length 1993;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

QY 121 ESLGLALGQAQEPHLSLLEAAEDLAQELLALRSLVELRALLQRPRTGSGPLELLSEALCS 180

Db 29 ESIGLALQAOEPLSHLEAEADLAQELALSLVELRALLOPRPGTSGPLELSEALCS 88
Qy 181 VRGPESTVPSLNWYEASDLMELVQEPESALPDSLSLSPACSELICALDSHPLSLLWRR 240
Db 89 VRGPESTVPSLNWYEASDLMELVQEPESALPDSLSLSPACSELICALDSHPLSLLWRR 148
Qy 241 LKPLILGKLLFAPDPTFFTRKMAQVNRTPFEELTLRLDRVREWEMLGPRIFTFMDSSNVA 300
Db 149 LKPLILGKLLFAPDPTFFTRKMAQVNRTPFEELTLRLDRVREWEMLGPRIFTFMDSSNVA 208
Qy 301 MLORLLOHODEGRRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTIGRVTEC 360
Db 209 MLORLLOHODEGRRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTIGRVTEC 268
Qy 361 LSLDKLEAAPSEAAALVSALQALLAEHRFWAGVVFGLGPDSSDPTHEPPTDLPQGHVRIKI 420
Db 269 LSLDKLEAAPSEAAALVSALQALLAEHRFWAGVVFGLGPDSSDPTHEPPTDLPQGHVRIKI 328
Qy 421 RMDIDVTRTKIRDPRWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRDPRWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 388
Qy 481 LYLOQMPYPCYVDDVFLRSLPLFLTLAWIYSVTITVKAIVREKETRLRDTWRAMGL 540
Db 389 LYLOQMPYPCYVDDVFLRSLPLFLTLAWIYSVTITVKAIVREKETRLRDTWRAMGL 448
Qy 541 SRAVLMLGWFLSCGLGPFLLSALLVLKIGDILPYSHPGVVFPLAFAFATVTSFLL 600
Db 449 SRAVLMLGWFLSCGLGPFLLSALLVLKIGDILPYSHPGVVFPLAFAFATVTSFLL 508
Qy 601 SAFFSRANLAAACGGLAYFSLYPLVYLCVAMRDLRDPAGGRVAASLLSPVAFGFCESIAL 660
Db 509 SAFFSRANLAAACGGLAYFSLYPLVYLCVAMRDLRDPAGGRVAASLLSPVAFGFCESIAL 568
Qy 661 LEEQEGQAWNVTGRTPTADVFLSLAQSGLLLDAAALYGLATWYLEAVCPQYGIPEPWN 720
Db 569 LEEQEGQAWNVTGRTPTADVFLSLAQSGLLLDAAALYGLATWYLEAVCPQYGIPEPWN 628
Qy 721 PFFRRSYWCGPRPPKSPACPTPLDPKVLVEAPPLGSPGVSRSLERKFPGPSOPALRG 780
Db 629 PFFRRSYWCGPRPPKSPACPTPLDPKVLVEAPPLGSPGVSRSLERKFPGPSOPALRG 688
Qy 781 LSLDPYQGHITAFLGHNGAGKTTTLLSILSGLFPSPGSAFILGHDDVRSSMAAIRPHLVC 840
Db 689 LSLDPYQGHITAFLGHNGAGKTTTLLSILSGLFPSPGSAFILGHDDVRSSMAAIRPHLVC 748
Qy 841 PQYNVLFDMLTVDHEHWFYGRILKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 900
Db 749 PQYNVLFDMLTVDHEHWFYGRILKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 808
Qy 901 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIWELLKLYREGRTILILSTHLLDEALLGD 960
Db 809 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIWELLKLYREGRTILILSTHLLDEALLGD 868
Qy 961 RVAVAGRLCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
Db 869 RVAVAGRLCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKKN 928
Qy 1021 GSQGSRVGTPOLLALVQHVPFGARLVEBELPHELVLVLPYTGADHGSFATLPRELDTRLAE 1080
Db 929 GSQGSRVGTPOLLALVQHVPFGARLVEBELPHELVLVLPYTGADHGSFATLPRELDTRLAE 988
Qy 1081 LRLTCYGISDTSLEIFLKVVEECAADTDMEDGSCGHLCTGIAGLDVTLRLKWPPOSTA 1140
Db 989 LRLTCYGISDTSLEIFLKVVEECAADTDMEDGSCGHLCTGIAGLDVTLRLKWPPOSTA 1048
Qy 1141 LENGEPAGSAPETDQSGPDVAVRGVQGWALTRQQLQALLKRLFLIARRSRGLFAQIVLP 1200
Db 1049 LENGEPAGSAPETDQSGPDVAVRGVQGWALTRQQLQALLKRLFLIARRSRGLFAQIVLP 1108
Qy 1201 ALFVGLALVFLSVLPFGHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGL 1260

Db 1109 ALFVGLALVFLSVLPFGHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGL 1168
Qy 1261 EEPVPQVSHSHRFSAPAEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPCCPAAAGP 1320
Db 1169 EEPVPQVSHSHRFSAPAEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPCCPAAAGP 1228
Qy 1321 PPOQAVTSGSVEVQNLTCGRNLSDFLVKTYPRLVROGLTKKVVNVRVYGGFSLGGRDPGL 1380
Db 1229 PPOQAVTSGSVEVQNLTCGRNLSDFLVKTYPRLVROGLTKKVVNVRVYGGFSLGGRDPGL 1288
Qy 1381 PSGQELGHSVEELWALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIWNKNGHSMWAF 1440
Db 1289 PSGQELGHSVEELWALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIWNKNGHSMWAF 1348
Qy 1441 VNRSASNLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI CVVFAM 1500
Db 1349 VNRSASNLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI CVVFAM 1408
Qy 1501 SFVPASPTLVLIIEERVTRAKHLQLMGGLSPTLYMGNFLMDCNLYLPACIIVLIFLAPQ 1560
Db 1409 SFVPASPTLVLIIEERVTRAKHLQLMGGLSPTLYMGNFLMDCNLYLPACIIVLIFLAPQ 1468
Qy 1561 QRAYVAPANI PAIILILILLYGWSITPLMYPASFPSPSTAYVLTICINLFIGINGSMAT 1620
Db 1469 QRAYVAPANI PAIILILILLYGWSITPLMYPASFPSPSTAYVLTICINLFIGINGSMAT 1528
Qy 1621 FVLELFSQKLOEVSRIKQVFLIPFPHFCIGRGLIDMVNRQAMADAFERLGRQFQSPUR 1680
Db 1529 FVLELFSQKLOEVSRIKQVFLIPFPHFCIGRGLIDMVNRQAMADAFERLGRQFQSPUR 1588
Qy 1681 WEVVGKLLAMVI QGPLFLLTLLQHRSQLLPQVRSLPLGGEDEDAVARERERVQG 1740
Db 1589 WEVVGKLLAMVI QGPLFLLTLLQHRSQLLPQVRSLPLGGEDEDAVARERERVQG 1648
Qy 1741 ATQGDVLVRLNLTKVYRGQRMVPAVDRICLIGIPPGECFGLLVNGAGKTTSTFRMVTGDTLA 1800
Db 1649 ATQGDVLVRLNLTKVYRGQRMVPAVDRICLIGIPPGECFGLLVNGAGKTTSTFRMVTGDTLA 1708
Qy 1801 SRGEAVLAGHSVAREPSSAAHLSMGYCPOSDAI FELLTQREHLELLARLGRVPEAQVAOTA 1860
Db 1709 SRGEAVLAGHSVAREPSSAAHLSMGYCPOSDAI FELLTQREHLELLARLGRVPEAQVAOTA 1768
Qy 1861 GSGLARLGLSWYADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMPSARRFLWN 1920
Db 1769 GSGLARLGLSWYADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMPSARRFLWN 1828
Qy 1921 SLLAVVREGSRVMLTSHSMEECEALCSRLATVWNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1829 SLLAVVREGSRVMLTSHSMEECEALCSRLATVWNGRFRCLGSPQHLKGRFAAGHTLTLRV 1876
Qy 1981 PAARSQPAAPAAAFVAAEFPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAEHGVDF 2040
Db 1877 ---RSQPAAPAAAFVAAEFPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAEHGVDF 1933
Qy 2041 SVSQTMLEEVLYFSKQGDDEDETEQKEAGVGDVPAPGLQHPKRVQSFLDDPSTAEVTL 2100
Db 1934 SVSQTMLEEVLYFSKQGDDEDETEQKEAGVGDVPAPGLQHPKRVQSFLDDPSTAEVTL 1993

RESULT 11
AAU04484
ID AAU04484 standard; protein; 1873 AA.
XX
AC AAU04484;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human PD-ATP-binding cassette (PD-ABC) protein form #2.
KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;

KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.
XX
OS Homo sapiens.
XX WO200153490-A1.
XX
XX 26-JUL-2001.
XX
XX 23-JAN-2001; 2001WO-US002191.
XX
XX 24-JAN-2000; 2000US-0177889P.
XX 30-JUN-2000; 2000US-0215405P.
XX (WARN) WARNER LAMBERT CO.
PA
XX
XX Johns MA, Tafuri SR, Wang M;
PI
XX
XX WPI; 2001-442259/47.
DR N-PSDB; AAS08707.
XX
XX New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
PT of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.
XX
XX Claim 10; Page 64-72; 77pp; English.
XX
XX The sequence represents human PD-ATP-binding cassette (PD-ABC) protein
CC form 2. PD-ABC maps to chromosome 19p13.3 and is expressed in various
CC tissues including spleen, thymus, peripheral blood leukocytes, bone
CC marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to
CC diagnose and treat cardiovascular disorders, inflammatory disorders,
CC dyslipidaemia, epilepsy, diseases related to abnormal calcium flux,
CC coronary artery disease, Tangier's disease, familial high-density
CC lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
CC insulin resistance, obesity, alcoholism, retinal degeneration,
CC hypertension and vascular disease. The sequences are also used in drug
CC screening assays
XX
XX Sequence 1873 AA;
SQ
Query Match 84.6%; Score 9213.5; DB 4; Length 1873;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1789; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 PPLEHHECHFPNKPPLPSAGTVPMLQGLICNVNNTCFPQLTPGEPGRSLNFNDLSRL 60
DB 47 PPLEHHECHFPNKPPLPSAGTVPMLQGLICNVNNTCFPQLTPGEPGRSLNFNDLSRL 106
QY 61 ADARTVLGGASAHRTLAGLKIATLRAARSTAQOPTKOSPPLMVDVAELLTSLRT 120
DB 107 ADARTVLGGASAHRTLAGLKIATLRAARSTAQOPTKOSPPLMVDVAELLTSLRT 166
QY 121 ESGLGALQOQEPHLSLEAEADLAQELLARSVELRALQPRGTSGPLELSEALCS 180
DB 167 ESGLGALQOQEPHLSLEAEADLAQELLARSVELRALQPRGTSGPLELSEALCS 226
QY 181 VRGSSSTVGPLNMYEASDLMELVGQEPESALPDSSLSACSELIGALDSHPISRLWRR 240
DB 227 VRGSSSTVGPLNMYEASDLMELVGQEPESALPDSSLSACSELIGALDSHPISRLWRR 286
QY 241 LKPLILGKLIAPDTPPTKRLMAQVNTFPELTLRDVREVWMLGPRIFTFNDSSNVA 300
DB 287 LKPLILGKLIAPDTPPTKRLMAQVNTFPELTLRDVREVWMLGPRIFTFNDSSNVA 346
QY 301 MLQRLQMQDEGRQPRPGRDHWEALRSFLDPSGGYSWQDAHADVGHVLGTLGRVTEC 360
DB 347 MLQRLQMQDEGRQPRPGRDHWEALRSFLDPSGGYSWQDAHADVGHVLGTLGRVTEC 406
QY 361 LSLDKLEAAPSEALVSRLQLLAHRFWAGVFLGPEDSSDPTHEPTPDLPQGHVRIKI 420
DB 407 LSLDKLEAAPSEALVSRLQLLAHRFWAGVFLGPEDSSDPTHEPTPDLPQGHVRIKI 466

QY 421 RMDIDVVTNRKIRDRFWDPCPAADPLTDLYRYVWGFFVYLQDLVERAAVVLGSANPRAG 480
DB 467 RMDIDVVTNRKIRDRFWDPCPAADPLTDLYRYVWGFFVYLQDLVERAAVVLGSANPRAG 526
QY 481 LYLQOMPYPICYDDVFLRLVLSRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRMAMGL 540
DB 527 LYLQOMPYPICYDDVFLRLVLSRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRMAMGL 586
QY 541 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVFLFLAAFAVATVTSFLL 600
DB 587 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVFLFLAAFAVATVTSFLL 646
QY 601 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVWRDLRAGRAAASLLSPVAFGFCESLAL 660
DB 647 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVWRDLRAGRAAASLLSPVAFGFCESLAL 706
QY 661 LEEQGEQAQWHNVGTRPTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 720
DB 707 LEEQGEQAQWHNVGTRPTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPMN 766
QY 721 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPOVSRSLEKRPFGSPQPALRG 780
DB 767 FPFRRSYWCGPRPKSPAPCPTPLDPKVLVEEAPPGLSPOVSRSLEKRPFGSPQPALRG 826
QY 781 LSLDFYQGHITAFILGHNGAGKTTTSLISGLPPSPGSAFILGHDVRSMAAIRPHLGVC 840
DB 827 LSLDFYQGHITAFILGHNGAGKTTTSLISGLPPSPGSAFILGHDVRSMAAIRPHLGVC 886
QY 841 PQYNVLFDMLTVDHEVWFYGRKLXGLSAAVVGPPQDRLLQDVLGVSQSVQTRHLSGMQR 900
DB 887 PQYNVLFDMLTVDHEVWFYGRKLXGLSAAVVGPPQDRLLQDVLGVSQSVQTRHLSGMQR 946
QY 901 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKTKREGRTLILSTHHLDEAELLGD 960
DB 947 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKTKREGRTLILSTHHLDEAELLGD 1006
QY 961 RVAVAVAGRLCCCGSPLFRRHLSGSGYLTLVKARLPLTTNEKADTMDEGSVDTROBKN 1020
DB 1007 RVAVAVAGRLCCCGSPLFRRHLSGSGYLTLVKARLPLTTNEKADTMDEGSVDTROBKN 1066
QY 1021 GSQGSRVGTPQLLALVQHWPVGARLVEELPHELVLPVLYTGAHSGSFATILFRELDTRLAE 1080
DB 1067 GSQGSRVGTPQLLALVQHWPVGARLVEELPHELVLPVLYTGAHSGSFATILFRELDTRLAE 1126
QY 1081 LRLTGIGISDTSLEELFLKVVECAADTMDEGSCQHLCGTAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGIGISDTSLEELFLKVVECAADTMDEGSCQHLCGTAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAAGAPETDQSGPDVAGRVQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1200
DB 1187 LENGEPAAGAPETDQSGPDVAGRVQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1246
QY 1201 ALFVGLALVFSLLVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDPGRARLEALLQBAGL 1260
DB 1247 ALFVGLALVFSLLVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDPGRARLEALLQBAGL 1306
QY 1261 EEPVVOHSHRFPADPEVAEVAIVKVLASGNWTPSPSPACQSPGARRLLPDCPAAAGGP 1320
DB 1307 EEPVVOHSHRFPADPEVAEVAIVKVLASGNWTPSPSPACQSPGARRLLPDCPAAAGGP 1366
QY 1321 PPPQAVTGSVEVQNTLGRNLSDFLVKTYPRVLQGLTKKWNVRYGGFSLGGRDPGL 1380
DB 1367 PPPQAVTGSVEVQNTLGRNLSDFLVKTYPRVLQGLTKKWNVRYGGFSLGGRDPGL 1426
QY 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLEDAQDSLKIFWNKGMHSMVAF 1440
DB 1427 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLEDAQDSLKIFWNKGMHSMVAF 1486
QY 1441 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNITKEQLSEALMASSVDVLSICVVFAM 1500
DB 1487 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNITKEQLSEALMASSVDVLSICVVFAM 1546
QY 1501 SPVPASFTLVLIERVTRAKHLQMGGLSPTLWLGNFLWDMCNLYLPACIVVLIFLAFQ 1560

Db 1547 SFVPASFTLVLEERVTRAKHLQMLGSLPTLVWLGFLDMCNVLPACIVLVFLAFQ 1606
 Qy 1561 ORAYVAPANLPALLILLYGHSITPLMYPASFFSPSTAYVLTINLFTGINGSMAT 1620
 Db 1607 ORAYVAPANLPALLILLYGHSITPLMYPASFFSPSTAYVLTINLFTGINGSMAT 1666
 Qy 1621 FVLELFSQKQEVSRILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1680
 Db 1667 FVLELFSQKQEVSRILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1726
 Qy 1681 MEVVGKULLAMVIOQPLFLFTLLQHSQQLPQVRSLPLGLGEEDVARERERVQ 1740
 Db 1727 MEVVGKULLAMVIOQPLFLFTLLQHSQQLPQVRSLPLGLGEEDVARERERVQ 1786
 Qy 1741 ATQGVVLNLTQVYRGQMPAVDRCLGIPGECFGLLGVNGAGKSTFRVMTGDTLA 1800
 Db 1787 ATQGVVLNLTQVYRGQMPAVDRCLGIPGECFGLLGVNGAGKSTFRVMTGDTLA 1844
 Qy 1801 SRGEAVLAGHSVAREPSSAAHLSMG-YCQSDAIFELL 1836
 Db 1845 SFTEHLLCIH-----HLLGTYCM---PIFVLL 1869

RESULT 12

ABG72695

ID ABG72695 standard; protein; 2167 AA.

XX ABG72695;

AC ABG72695;

XX ABG72695;

DT 10-MAR-2003 (first entry)

XX Mouse ATP-binding cassette transporter-like protein, ABCI.

DE Mouse; ATP-binding cassette transporter-like protein; ABCI;

KW lipid transport; cardiovascular disease; hypertriglyceridaemia;

KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;

KW nervous system disorder; Stargardt disease; degenerative disorder;

KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;

KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;

KW acquired immunodeficiency disorder; leukaemia; neutropenia; anaemia;

KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;

KW hypothyroidism; obesity; diabetes; reproductive disorder;

KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;

KW autoimmune disease; inflammatory disease; multiple sclerosis.

XX Mus musculus.

OS Mus musculus.

XX Key

XX Location/Qualifiers.

XX Peptide

XX 1..46

XX /label= Signal_peptide

XX Protein

XX 47..2167

XX /label= Mature_ABCI

XX US2002127647-A1.

XX 12-SEP-2002.

XX 28-NOV-2001; 2001US-00995542.

XX 28-NOV-2000; 2000US-0253520P.

XX (SHUT/) SHUTTER J.

XX (ULIA/) ULIAS L.

XX Shutter J, Ulia L;

XX WPI; 2003-147394/14.

XX N-PSDB; ABX14665.

XX Novel ATP-binding cassette transporter-like polypeptides and

XX polynucleotides useful for diagnosing, preventing, treating disorders

XX involving immune, nervous system, thyroid, hypothalamus and impaired

PT transport of lipids.
 XX Claim 13; Fig 1; 149pp; English.
 PS The invention relates to an isolated murine and human ATP-binding
 CC cassette transporter-like (ABCI) polypeptide, or the amino acid sequence
 CC encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
 CC 3111. Also included are the nucleic acids encoding the ABCI proteins, or
 CC vectors, host cells, ABCI binding agents, a selective binding agent or
 CC its fragment comprising at least one complementarity determining region
 CC (CDR) with specificity for ABCI which (produced by immunising an animal
 CC with ABCI), a hybridoma that produces the CDR, viral vectors, an ABCI
 CC fusion polypeptide, a device comprising a membrane suitable for
 CC implantation (permeable to the protein and impermeable to materials
 CC detrimental to the cells, and cells encapsulated within the membrane)
 CC where the cells secrete ABCI, an ABCI transgenic non-human mammal and an
 CC array of ABCI nucleic acid molecules. The ABCI polypeptide, nucleic acids
 CC and modulators are useful for the diagnosis and/or treatment of diseases
 CC and conditions involving impaired transport of lipids, including
 CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
 CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
 CC involving functional and trophic disturbances of the nervous system such
 CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
 CC fibrosis, conditions involving multidrug resistance, conditions involving
 CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
 CC neutropenia, anaemia and autoimmune diseases, conditions involving the
 CC thyroid e.g. hyper and hypothyroidism; conditions involving the
 CC hypothalamus including obesity, diabetes, reproductive disorders, energy
 CC balance disorders, peripheral neuropathies including myelinopathies and
 CC axonopathies, autoimmune and inflammatory diseases involving the nervous
 CC system including multiple sclerosis. The present sequence represents
 CC murine ABCI

Sequence 2167 AA;

Query Match 77.0%; Score 8389.5; DB 6; Length 2167;
 Best Local Similarity 76.3%; Pred. No. 0;
 Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;
 Qy 1 PPLEHCHFPNKLPSAGTVPMIQLGICNVNNTCFPQITGEEPRGLSNFNDSIVSRLL 60
 Db 47 PPLEHCHFPNKLPSAGTVPMIQLGICNVNNTCFPQITGEEPRGLSNFNDSIVSRLL 106
 Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARATQ-----OPTKQSPLEPPMLDVALLTS 116
 Db 107 ADTRTVLGGHSIQMDLDALGKLIPLVRAVGGARQPSDQPTSQ-----SVTKLLEK 159
 Qy 117 LLRTESLGLAQOQEPHLSLEAAEDLAQELLALRLSLVELRALRLQPRGTSGPLELISE 176
 Db 160 ILQASLDPLVGOAQDSMRKESDAIRDLAQELLTLPMLRALRLRRPGSGSLVSE 219
 Qy 177 ALCSVRGSSVTGPSLNWYEASDLMLVGPESALPDSSISGACSEIIGALDSHPLSRL 236
 Db 220 ALCSVTGSPSPGSLNWEANQNFMPGVPAPALPDNSLSPACSEFVGLTDDHPVSRLL 279
 Qy 237 LWRRLKPLILGLKLPADPTPTKRLMAQVNTFFELTLRLDVRVWEMGLPRIFPMDS 296
 Db 280 LWRRLKPLILGLKLPADPTPTKRLMAQVNTFFELTLRLDVRVWEMGLPRIFPMDS 339
 Qy 297 SNVAMLQRLQOMQDEGRPRPGGRDHMEALRSFLDPGSGGYSDAHADVGLHVTGLCR 356
 Db 340 TNVAMLQRLDVGVTGQGTQPTPRAQKLEALIKDFDPSRGYSWREAHADMGRLAGILQ 399
 Qy 357 VTECLSDKLEAPSEALVSRLQLLAHRFWAGVFLGPESSDPTTEHPTPDIGPGHV 416
 Db 400 MMECVSLDKLEAVPSEALVSRLQLLAHRFWAGVFLGPESSDPTTEHPTPDIGPGHV 459
 Qy 417 RIKIRMDIVVTRTNKIRDRFWDGPAADPLTDLRYVWGGVYIQLDLVERAARVLSGAN 476
 Db 460 RFIKIRMDIVVTRTNKIRDRFWDGPAADPLTDLRYVWGGVYIQLDLVERAARVLSGAN 519
 Qy 477 PRAGLYLQOMPYPCYVDVFLRVLRSRLPLFLTLTAWYSVTLTKAVVREKETRLDRMT 536

Db 520 SRTGLYLQMPHPCYVDDVFLRVLRSRLPLFLTLAMIYSVALTAKVAVREKRETLRETMR 579
Qy 537 AMGLSRVILWLCWFLSCIGLPPFLSAALLVLVLKGDILPYSHPGVWFLFLAAFAVATVQ 596
Db 580 AMGLSRVILWLCWFLSCIGLPPFLSAALLVLVLKGNILPYSHPVVIFLFLAAFAVATVAQ 639
Qy 597 SFLLSAFSFRANLAAACGGLAYFSLPYLYLVCVAMRDLRLLPAGGRVAAASLLSPVAFGCE 656
Db 640 SFLLSAFSFRANLAAACGGLAYFALYLYLVCVAMRDLRLLHGLLAASLLSPVAFGCE 699
Qy 657 SLALLEQGEQAQHNVTGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPQGYGIP 716
Db 700 SLALLEQGEQAQHNVTGTRPTADVFLSAQVSAFLLDVAVIYGLWYLEAVCPQGYGIP 759
Qy 717 EPNWFPFRSRYWCOPRPPKSPAPCPPTDPLDKVLYERAPPGLSPGVSVRSLEKPPGSPQP 776
Db 760 EPNWFPFRSRYWCOPRPPKSSVLPAPQDPKVLVEEPPLGLVPGVSRIGLKHFRGCPQP 819
Qy 777 ALRGLSLDFYQCHITAFLGHNGAGKTTTSLTSLGLFPPSGSAPFILGHVDRSSMAAIRPH 836
Db 820 ALQGLNLDYFEGHITAFLGHNGAGKTTTSLTSLGLFPPSGSASILGHVQTNMAAIRPH 879
Qy 837 LGVCPQYNVLFDMULTVDEHVMFYGLKGLSAAVGPBQDRLLQDVGIVSKQSVQTRHLSG 896
Db 880 LGICPQYNVLFDMULTVEHVMFYGLKGLSAAVGPBQDRLLQDVGIVSKQSVQTRHLSG 939
Qy 897 GMQRKLSVAFAVGGSOVWILDEPTAGVDPASRRGIWELLKYREGTTLILSTHLLDEAE 956
Db 940 GMQRKLSVAFAVGGSRVWIMDEPTAGVDPASRRGIWELLKYREGTTLILSTHLLDEAE 999
Qy 957 LLGDRVAVAGRLCCCGSLPLFRHHGSGYLLTVKARLPLTNE-KAPDMEGSDVTR 1015
Db 1000 LLGDRVAVAGRLCCCGSLPLFRHHGSGYLLTVKARLPLTNE-KAPDMEGSDVTR 1054
Qy 1016 QEKNGSQGS-----RVGTQQLALVQHWVPGARLVEELP 1050
Db 1055 REKSDCNGRTSDTAFTRGTSKSNQAPAGVPIPTSTARILELVQHVPGAQVEDLP 1114
Qy 1051 HELVLVLPYTGADHGSFATLPRELDTLRLAEURLTGYGSDTSLEELFLKVEECAADTM 1110
Db 1115 HELVLVLPYTGADHGSFATLPRELDTLRLAEURLTGYGSDTSLEELFLKVEECAADTM 1172
Qy 1111 EDGSCGCHLTCTAGLDVTLRLKMPPOETALENGEPAGSA--PETDQSGPDVAVRQGV 1168
Db 1173 GDSRPQHURT-----CTQPTGPEASVLENGELAKVLDPQAPQGLAPNA-AQVQGV 1225
Qy 1169 ALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1228
Db 1226 TLTCQQLRALLHKEFLARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1285
Qy 1229 MYGQVQFFSDEADPGDGRARLLBALLOEAGLEBPVVOHSHR-----FSAPEV 1277
Db 1286 MYGQVQFFSDEADPGDGRARLLBALLOEAGLEBPVVOHSHR-----FSAPEV 1345
Qy 1278 PAEVAKLASGNWTPESPACQCSQPGARRLLPDCPAAAGPPPPQAVTSGSRVQNL 1337
Db 1346 PPDVASILASGNWTPESPACQCSQPGARRLLPDCPAGAGPPPPQAVTSGSRVQNL 1405
Qy 1338 GRNLSDFLVKTYPLRVQGLTKKWNNEVRYGFSGLGRDPGLPSGQELGRSVEELWALL 1397
Db 1406 GRNLSDFLVKTYPLRVQGLTKKWNNEVRYGFSGLGRDPGLPSGQELGRSVEELWALL 1465
Qy 1398 SPLPGGALDRLVKNLTAHWSLDAQDSLTKWFKGHWMSVAFVNRASNAILRAHLPPGP 1457
Db 1466 SPOGNAIDRLNLTQWALGDARNSLTKWFKGHWMSVAFVNRANGLLHALLPSGP 1525
Qy 1458 ARHAHSITTLNHPUNLTKQLESEALMASSVDVLVSCVVFAMSFVPASFTLVLIBERT 1517
Db 1526 VRHAHSITTLNHPUNLTKQLESEATLASSVDVLVSCVVFAMSFVPASFTLVLIBERT 1585
Qy 1518 RAKHLQMGGLSPTYLWGLNFWLDMCNLYLPACITVLI FLAQORAYVAPANLALLL 1577
Db 1586 RAKHLQVSGLPQTLWGLNFWLDMCNLYLVAVCIWVFLFLAQORAYVAPANLALLL 1645

Qy 1578 LLYGWSITPLMYPASRPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEVSRI 1637
Db 1646 LLYGWSITPLMYPASRPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEVSRI 1705
Qy 1638 LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPFLRWVGVKGNLLAMVIOGPL 1697
Db 1706 LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPFLRWVGVKGNLLAMVIOGPL 1765
Qy 1698 FLITLTLLOHRSOLLQPPVRSLPLGEDEEDVARERVRVQATQGDVLVRLNLTQVYR 1757
Db 1766 FLITLTLLOHRSOLLQPPVRSLPLGEDEEDVARERVRVQATQGDVLVRLNLTQVYR 1825
Qy 1758 GORMPAVDRCLGIPGCECFGLLVGNGAGKSTFRWVTGDTLASRGAVALAGHSVAREPS 1817
Db 1826 GORMPAVDRCLGIPGCECFGLLVGNGAGKSTFRWVTGDTLASRGAVALAGHSVAREPS 1885
Qy 1818 AAHLGMYCQPSQDAIFELITGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYADRP 1877
Db 1886 AAHLGMYCQPSQDAIFELITGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYADRP 1945
Qy 1878 GTYSGGNKRLATALVGDPAVVFLEDEPTTGMDDPSARRFLMNSLLAVVREGSRVMTSH 1937
Db 1946 GTYSGGNKRLATALVGDPAVVFLEDEPTTGMDDPSARRFLMNSLLAVVREGSRVMTSH 2005
Qy 1938 SMEECALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLTLRVPAAARSQAAAFVAAEFP 1997
Db 2006 SMEECALCSRLAIMVNGRPRCLGSPQHLKGRFAAGHTLTLRVPAAARSQAAAFVAAEFP 2065
Qy 1998 GSELRAHGRLRFQPLPGGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYFSKD 2057
Db 2066 GSELRAHGRLRFQPLPGGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYFSKD 2125
Qy 2058 QGKDEPTEQKEAGVVDPAQGLQHPKRVSQFLDDPSTAEVL 2100
Db 2126 QGKDEPTEQKEAGVVDPAQGLQHPKRVSQFLDDPSTAEVL 2167
RESULT 13
ABG72697
ID ABG72697 standard; protein; 1550 AA.
XX
AC ABG72697;
XX
DT 10-MAR-2003 (first entry)
XX
DE Human ATP-binding cassette transporter-like protein, ABCL1550.
XX
KW Human; ATP-binding cassette transporter-like protein; ABCL; ABCL1550;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothyroidism; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
OS Homo sapiens.
XX
XX US2002127647-A1.
XX
PD 12-SEP-2002.
XX
XX 28-NOV-2001; 2001US-00995542.
XX
XX 28-NOV-2000; 2000US-0253520P.
XX
XX (SHUT/) SHUTTER J.
XX (ULIA/) ULIA L.
XX

PI Shutter J, Ulias L;
 XX WPI; 2003-147394/14.
 DR N-PSDB; ABX14667.
 XX
 XX
 PT Novel ATP-binding cassette transporter-like polypeptides and
 PT polynucleotides useful for diagnosing, preventing, treating disorders
 PT involving immune, nervous system, thyroid, hypothalamus and impaired
 PT transport of lipids.
 XX
 XX
 PS Claim 13; Fig 3; 149pp; English.
 XX
 CC The invention relates to an isolated murine and human ATP-binding
 CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
 CC encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
 CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
 CC vectors, host cells, ABCL binding agents, a selective binding agent or
 CC its fragment comprising at least one complementarity determining region
 CC (CDR) with specificity for ABCL which (produced by immunising an animal
 CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
 CC fusion polypeptide, a device comprising a membrane suitable for
 CC implantation (permeable to the protein and impermeable to materials
 CC detrimental to the cells, and cells encapsulated within the membrane)
 CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
 CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
 CC and modulators are useful for the diagnosis and/or treatment of diseases
 CC and conditions involving impaired transport of lipids, including
 CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
 CC hypercholesterolaemia, tangier disease, dyslipidaemias, conditions
 CC involving functional and trophic disturbances of the nervous system such
 CC as Stargard disease, degenerative and inflammatory retinopathy, cystic
 CC fibrosis, conditions involving multidrug resistance, conditions involving
 CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
 CC neutropenia, anaemia and autoimmune diseases, conditions involving the
 CC thyroid e.g. hyper and hypothyroidism; conditions involving the
 CC hypothalamus including obesity, diabetes, reproductive disorders, energy
 CC balance disorders, peripheral neuropathies including myelinopathies and
 CC axonopathies, autoimmune and inflammatory diseases involving the nervous
 CC system including multiple sclerosis. The present sequence represents
 CC human ABCL truncated variant, ABCL1550
 XX
 XX Sequence 1550 AA;
 SQ
 Query Match 69.7%; Score 7596; DB 6; Length 1550;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 121 ESLGALGQAQBPPLHSLLEAEDLAQELLALSLVELRALLQRPRTSGPLELSEALCS 180
 DB 29 ESLGALGQAQBPPLHSLLEAEDLAQELLALSLVELRALLQRPRTSGPLELSEALCS 88
 QY 181 VRGSPSTVGPLNWTYEASDLMELVQEPESALPDSSLPACSELIGALDHPHLLWRR 240
 DB 89 VRGSPSTVGPLNWTYEASDLMELVQEPESALPDSSLPACSELIGALDHPHLLWRR 148
 QY 241 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEETLLRDVREVMELGPRFTFMDNSSVA 300
 DB 149 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEETLLRDVREVMELGPRFTFMDNSSVA 208
 QY 301 MLQRLQMDQEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTLLGRVTEC 360
 DB 209 MLQRLQMDQEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTLLGRVTEC 268
 QY 361 LSLDKLEAPSEALVSRALQLLAERHFWAGVVFGLPEDSSDPTBHPDLPQGHVRIKI 420
 DB 269 LSLDKLEAPSEALVSRALQLLAERHFWAGVVFGLPEDSSDPTBHPDLPQGHVRIKI 328
 QY 421 RMDIDVTRTKIRDFWDPGPAADPLTLDRVWVGFFVLODLVERAAVRVLSGANPRAG 480
 DB 329 RMDIDVTRTKIRDFWDPGPAADPLTLDRVWVGFFVLODLVERAAVRVLSGANPRAG 388
 QY 481 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKA VVREKETRLDRTMAMGL 540
 DB
 389 LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKA VVREKETRLDRTMAMGL 448
 541 SRAVLWLGWFLSCLGPFLLSAALLVLVLKGLDILPYSHPGVVPFLFLAAFAVATVTSQFLL 600
 449 SRAVLWLGWFLSCLGPFLLSAALLVLVLKGLDILPYSHPGVVPFLFLAAFAVATVTSQFLL 508
 601 SAFFSRANLAAACGGLAYFSLYLYVVLVWMDRULPAGGRVAASLLSPVAFGCGCSLAL 660
 509 SAFFSRANLAAACGGLAYFSLYLYVVLVWMDRULPAGGRVAASLLSPVAFGCGCSLAL 568
 661 LEEQGEQAQHNVTGTRPTADVPSLAQVSGLLLDLAALYGLATWYLA VCPGQYGIPEPMN 720
 569 LEEQGEQAQHNVTGTRPTADVPSLAQVSGLLLDLAALYGLATWYLA VCPGQYGIPEPMN 628
 721 FPFRRSYWCGPPPKSPAPCPPTPLDPKVLVEAPGLSPGVSVRSLEKFPSPGQPALRG 780
 629 FPFRRSYWCGPPPKSPAPCPPTPLDPKVLVEAPGLSPGVSVRSLEKFPSPGQPALRG 688
 781 LSLDFYQGHITAFLGHNGAGKTTTLLSLSGLFPSPGGSFALICH DVRSMAAIRPHLVC 840
 689 LSLDFYQGHITAFLGHNGAGKTTTLLSLSGLFPSPGGSFALICH DVRSMAAIRPHLVC 748
 841 POYNVLFDMLTVDHWFYGRUKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGMQR 900
 749 POYNVLFDMLTVDHWFYGRUKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGMQR 808
 901 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLLSTHHLDEALLGD 960
 809 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLLSTHHLDEALLGD 868
 961 RVAVVAGRLCCCGSPLFRRHLGSGYYLTVKARLPLTTNEKADTDMEGSVDTREKKN 1020
 869 RVAVVAGRLCCCGSPLFRRHLGSGYYLTVKARLPLTTNEKADTDMEGSVDTREKKN 928
 1021 SGQSRVGTQQLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTRLAE 1080
 929 SGQSRVGTQQLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTRLAE 988
 1081 LRLTCYGISDTSLEEIFLKVVEECAADTDMEGSCGQHLCTGIA GLDVTLRKMPQETA 1140
 989 LRLTCYGISDTSLEEIFLKVVEECAADTDMEGSCGQHLCTGIA GLDVTLRKMPQETA 1048
 1141 LENGEPAGSAPETDQSGPDVGRVQGWALTQQLQALLKRLFLARRRRLGFAQIVLP 1200
 1049 LENGEPAGSAPETDQSGPDVGRVQGWALTQQLQALLKRLFLARRRRLGFAQIVLP 1108
 1201 ALFVGLALVFSLIVPPFGHYPALRLSPTMYGAVSFSEDAFGDPGRARLLLEALLQEA 1260
 1109 ALFVGLALVFSLIVPPFGHYPALRLSPTMYGAVSFSEDAFGDPGRARLLLEALLQEA 1168
 1261 BEPPVQSHSRFSAPEVPAEVAKVULASGNWTPESPACQSPGARRLLPDCPAAAGP 1320
 1169 BEPPVQSHSRFSAPEVPAEVAKVULASGNWTPESPACQSPGARRLLPDCPAAAGP 1228
 1321 PPOAVTSGSEVQNLTCRNLSDFLVKTYPRLVROGLTKKWVNEVRVGGPSLGGDRDGL 1380
 1229 PPOAVTSGSEVQNLTCRNLSDFLVKTYPRLVROGLTKKWVNEVRVGGPSLGGDRDGL 1288
 1381 PSGQELGRSVEELWALLSPLFCGALDRVLKNTLAWAHS LDAQDSLKIWNKNGHSMVAF 1440
 1289 PSGQELGRSVEELWALLSPLFCGALDRVLKNTLAWAHS LDAQDSLKIWNKNGHSMVAF 1348
 1441 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVPAM 1500
 1349 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVPAM 1408
 1501 SFVPASFTLVLIERVTRAKHLQLMGGLSPTLYMGNFLWDMCNVLPACI VVLI FLAQ 1560
 1409 SFVPASFTLVLIERVTRAKHLQLMGGLSPTLYMGNFLWDMCNVLPACI VVLI FLAQ 1468
 1561 QRAYVAPANLPA LLLLLLLLYGWSITPL 1587
 1469 QRAYVAPANLPA LLLLLLLLYGWSITPL 1495

RESULT 14

AAB38111
 ID AAB38111 standard; protein; 2261 AA.
 AC AAB38111;
 DT 29-JAN-2001 (first entry)
 XX Human ABC1 cholesterol transporter mutant, V771M.
 DE
 XX Human ABC1 cholesterol transporter; chromosome 9q31;
 KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
 KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
 KW cardiovascular disease; coronary artery disease; coronary restenosis;
 KW cerebrovascular disease; peripheral vascular disease;
 KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
 KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
 KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
 KW mutant.
 XX Homo sapiens.
 OS
 XX WC200055318-A2.
 PN
 XX 21-SEP-2000.
 PD
 XX 15-MAR-2000; 2000WO-IB000532.
 XX
 XX 15-MAR-1999; 99US-0124702P.
 PR 08-JUN-1999; 99US-0138048P.
 PR 17-JUN-1999; 99US-0139600P.
 PR 01-SEP-1999; 99US-0151977P.
 XX
 XX (UYBR-) UNIV BRITISH COLUMBIA.
 PA (XENO-) XENON BIORSEARCH INC.
 PA
 XX Hayden MR, Wilson AR, Pimstone SN;
 PI WPI; 2000-587528/55.
 XX
 XX New ABC1 polypeptide is useful for treating diseases associated with ABC1
 PT biological activity, e.g. Alzheimer's disease, Huntington's disease and
 PT cancer.
 PT
 XX Example; Page; 229pp; English.
 PS
 XX The invention relates to the human ABC1 cholesterol transporter protein
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary restenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds. It
 CC further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or

CC prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary restenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
 CC present sequence represents a mutant human ABC1 cholesterol transporter
 CC associated with an altered cholesterol level and therefore an altered
 CC risk of cardiovascular disease. Note: The present sequence is not shown
 CC in the specification, but is derived from the native human ABC1 shown on
 CC pages 152-157
 XX Sequence 2261 AA;
 SQ
 Query Match 51.5%; Score 5611.5; DB 3; Length 2261;
 Best Local Similarity 50.0%; Pred. No. 0;
 Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
 QY 1 PPLEHCHFPNKPPLPSAGTVPMLQGLICNVNTCTPQTPGEGPGLSNFNDLSYRL 60
 DB 47 PPYEQECHFPNKPPLPSAGTLPWVQGIICNANNPCFRYPTEGAPGVGVGNFKSIVARLF 106
 QY 61 ADARTVLGGASAHRTLAGLGLKLIATLRAARSTAQ-----PQ 96
 DB 107 SDARLLLSQKDTSMKMRKVLTQQIKSSNLKLDLVYDNETFSFLYHNSLPK 166
 QY 97 PTKQSPLEPML-----DVAEL----- 113
 DB 167 STVDKMLRADVILHKVFLQYQLHTSLCNGSKSEMIQLGDQEVSELCLPREKLA 226
 QY 114 -----LTSLLRT--ESLGLALGQAQEPHLSLEAEADLAQELLARSLSVELR--A 159
 DB 227 RVLRSNMDILKPIRLTINSTPPPSKELABATKTLHSLGTLAQELFSMSMDMREVM 286
 QY 160 LLQRPRTSGLEL---LSEALCSVRGSPSTVGPLNWEASDLMELVG---QEPESAL 212
 DB 287 FLTNVSSSSTQIYQAVSRIVCGHPEGGLKTKSLNWNEDNYYKALFGNGTEEDATF 346
 QY 213 PDSLSLPACSELIGALDSDHPLSLRRLKPLILGLLFPADPTPTFRKLMQAVNRTEEL 272
 DB 347 YDNSTTTPYCNLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATQVMAEVNKTPOEL 406
 QY 273 TLIRDVREVMELGPRIFTFMDNDSNVAMLQRLQWQDEGR-RQPRGGD----- 322
 DB 407 AVFHDLEGWMEELSPKIWTFMENSQEMDLVRMLDSRDNDFHWEQQLDGLDWTADIVAF 466
 QY 323 ---HMEALRFLDPGSGG--YSWQDAHADVGHVLVGTILGRVTECLSLDLKLEAAPSEAA 377
 DB 467 LAKHPEDVQS-----SNGSVYTWREAFNEINQAIRISRFWECVNLNLEPIATEVW 521
 QY 378 RALQLLAHRFWAGVVFGLPEDSDPTHTPDLPGLGHVRIKIRMDIDVVYTRNKIKDRF 437
 DB 522 KSMELDERKFWAGIVFTGTPGSIELPH-----HVKYKIRMDIDNVERTNKIKDGY 573
 QY 438 WDFGPAADPTLDIYVWGGFVYLQDLVERAAVLSGANPRAGLYLQOMVPCVVDVDEL 497
 DB 574 WDFGPRADPPEDRYRYWGGSPAYIQDVVVEQAIIRVLTGTEKTKGVYMQMPPCYVD 633
 QY 498 RVLRSRSLPLTLAWIYSVTLVKAVVRKEKTRLDTRAMGLSRVAVLWGLFSLCLGPF 557
 DB 634 RVNRSRNPLEWTAWIYSVAIVKIGIVYEKARLKTMRIMGLDNLNLSWFISSLPL 693
 QY 558 LLSAALLVLVLKGLDILPYSHPGVVFVFLAAFAVAVTQSFLLSAFPSRANLAACGLA 617
 DB 694 LVSAGLLVWLKGLNLLPYSDPSVVFVFLSVFAVVTILQCLFLISTLSFRANLAACG 753
 QY 618 YFSILYPLVYLCVAVRDELPAGRVAASLLSPVAFGCGESIALLEQEGAGQHWGTRP 677
 DB 754 YFTLYPLVYLCVAVQDMYMGFTLKIFASLLSPVAFGCGEYFALFEGIGVQWMDNLF 813
 QY 678 T-ADVFLAQVSGLLILLDAALYGLATWYLEAVCPGQYGIPEPWNPFFRRSYWCGPRPKS 736

CC The invention relates to the human ABC1 cholesterol transporter protein
 CC (B38032) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary stenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds. It
 CC further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or
 CC prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary stenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, X-linked adrenoleukodystrophy and cancer.
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of Genbank Accession No: CAA1005.1 and X75926, and the nucleic
 CC acid with the exact sequence as Genbank Accession No: AJ012376.1. The
 CC present sequence represents a mutant human ABC1 cholesterol transporter
 CC associated with an altered cholesterol level and therefore an altered
 CC risk of cardiovascular disease. Note: The present sequence is not shown
 CC in the specification, but is derived from the native human ABC1 shown on
 CC pages 152-157

XX Sequence 2261 AA;

Query Match 51.5%; Score 5611.5; DB 3; Length 2261;
 Best Local Similarity 50.0%; Pred. No. 0;
 Matches 1116; Conservative 361; Mismatches 596; Indels 161; Gaps 26;

Qy 1 PPLEHCHFPNKPPLSAGTVPWVQGLICNVNNTFCFPLTPGEPGRLSNFNSLSRL 60
 Db 47 PVEQHECHFPNKPPLSAGTVPWVQGLICNVNNTFCFPLTPGEPGRLSNFNSLSRL 106
 Qy 61 ADARTVGGASAHRTLAGLKLIALTAARSTAG-----PQ 96
 Db 107 SDARLLLYSOKOTSMKMRKVLRTLOQIKSSSNLKLQDFLDVNETFSGFLYHNLSPK 166
 Qy 97 PTKQSPLEPPML-----DVAEL-----113
 Db 167 STVDKMLRADVLHKVFLQYQVHLTLTLGKSGSKSEMIQLGDQVSELCGLPREKLAEE 226
 Qy 114 -----LTSLLRT--ESLGLAQOAEPLHSLLEAEADLAQELLALRLSLVELR---A 159
 Db 227 RVLRSNMDILKPIRLTLNSTSPPSKELAEATKTLHSLGLTAQELFSMRSDMRQVM 286
 Qy 160 LLQRPRTSGPLEL---LSEALCVSPGSPSTVSGPSLWYASDLMLVG-----QEPESAL 212
 Db 287 FLTNVNSSSSTQIQVARSIVCGHPGEGGLIKLSLWYEDNNYKALFGNGTEDEATF 346
 Qy 213 PDSLSLSPACSELICALDHPRLSLRLWRLKPLIILGKLLFAPDPTFRKLMAQVNRTEEL 272
 Db 347 YDNTTTCYNDLMKNLSESSPLSRIIWKALPLIVGKILYTPDTPATQVMAEVNKTFOEL 406
 Qy 273 TLLRDFREVWMLGRIFTFMNDSSNVMAQLRLQLOMQDEGR--RQRPFGGRD-----322
 Db 407 AVFHDLEGMWBEELSPKTIWTFMENSEQEMDLVRLMLDSRDNHDFWQQLDGLDWTQAIVAF 466

Qy 323 ---HMEALRSFLDPGSG--YSQDAHADVGHVGLVGTGRVTECLSLDKLEAAPEAAALVS 377
 Db 467 LAKHPEDVQS-----SNGSVYTWREAFNEFNQAIRISRFMECNLKNLEPIATEWFLIN 521
 Qy 378 RALQLLAEHRFWAGVVFPLGPESSDTEHPTDPLGPHGVRKIRMDIDVVYTRTNKIDRF 437
 Db 522 KSMELDERKFWAGIVFTGTPGSIELPH-----HVYKIRMDIDVNERTNKIDGY 573
 Qy 438 WDPGPAADPTUDRYVWGVVYQDLVERAAVVLGNSANPRAGLYLOOMPYPCVVDVFL 497
 Db 574 WDPGPAADPTUDRYVWGVVYQDLVERAAVVLGNSANPRAGLYLOOMPYPCVVDVFL 633
 Qy 498 RVLRSRSLPLTLTAWYSVTLTKAVVREKETRLRDTMRAMGLSRVAVLWGLFSLCLGPF 557
 Db 634 RVMSRSMPLFTLWYISVAVIIGVYKEAKLKTMRINGLNDNSLWFSWISSLIPL 693
 Qy 558 LLSAALLLVVLKGLDILPYSHPGVWFLFAAFVAVTQSFLLSAFSSRANLAACGGLA 617
 Db 694 LVSAGLLVVLKGLDILPYSDPSVVFVFLSVFAVVTILQCLFLISTLFSRANLAACGGII 753
 Qy 618 YFSLYLPYVLGVAMRDLRPLAGGRVAAVLLSPVAFGFCESLALLEBQEGEAQHNVTRP 677
 Db 754 YFTLYLPYVLGVAMRDLRPLAGGRVAAVLLSPVAFGFCESLALLEBQEGEAQHNVTRP 813
 Qy 678 T-ADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOVGIPEPWNFPRRSYWCGRPPKS 736
 Db 814 VEEDGNLTTSVSMMLFDTFLYGVMTWYIEAVPQGIIPRPWYFPTCKYIWFGESEDEK 873
 Qy 737 PAPCPPTPLD--PKVLVEEAPGLSPGVSRLSEKRFPGSPQALRGLSLDFYQGHITAFGL 795
 Db 874 SHFGSNQKRISCEIMEEETHLKGVSIGNLVKYRDKMVAVDGLALNFVEGQITSLFLG 933
 Qy 796 HNGAGKTTLSLUSGLFPPSGGSAFILGHVRSMAIRPHLGVCPYVNVFLDMLTVDEH 855
 Db 934 HNGAGKTTLSLUSGLFPPSGGSAFILGHVRSMAIRPHLGVCPYVNVFLDMLTVDEH 993
 Qy 856 VWFVGRKLGLSAAVVGPEQDRLLQDVGL--VSKQSVOTRHLSSGMQKLSVAIAFVGGSQV 914
 Db 994 IWFYARLKLSEKHVKAEMQMALDVGLPSSKLKSTLSQSGGQKLSVALAFVGGSKV 1053
 Qy 915 VILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 974
 Db 1054 VILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 1113
 Qy 975 SPLFLRRHLGSGYILVILKARLPLTNN-----EKADTMEGSDVTRQEKNGS 1022
 Db 1114 SSLFLKNQLGTGYLTLVKKDVESLSSCRNSSSTVYLKKBDSVSSQSSDAGLSGDHS 1173
 Qy 1023 QGSRVGTGPOLLALVOHWPVGLARLVEELPHELVLVLPYTGADHGSFATLFLRELDTLAEALR 1082
 Db 1174 DTUTIDVSAISNLIKRVSEARLVEDIGHELTVLPYEAKEGAFVLEFHEIDRLSDLG 1233
 Qy 1083 LTGYGISDTSLEBIFLKVVEECAADTMDMEDGSCQHLCTGIAGLDVTLRLKMPPEQTALE 1142
 Db 1234 ISSYGISSETLEBIFLKVVEECAADTMDMEDGSCQHLCTGIAGLDVTLRLKMPPEQTALE 1292
 Qy 1143 --NGEPAGSAPETDQSGDPDVG--RVQGWALTFRQLOALLLKRFLLARRSRGLFAQIV 1198
 Db 1293 PNDSDIDPESRETDLLSGMDGKGSYQVKGWKLTKQQFVALLWKRLLIARRSRGFFAQIV 1352
 Qy 1199 LPALFVGLALVFLSIVPPFGHYPALRLSPYMGAGVQSFESDAPGDPGRARLLLEALQEA 1258
 Db 1353 LPAVFVCIALVFLSIVPPFGHYPALRLSPYMGAGVQSFESDAPGDPGRARLLLEALQEA 1412
 Qy 1259 G-----LBEPPVQHSRHSFAPEVPAEVAKVLASGNWTPSPSPACQSQGARR 1308
 Db 1413 GFGTRCMEGNPIDPTPCQAGEEWTAPVPTQIMDLFQNGNWTWQNSPACQSSDKIKK 1472
 Qy 1309 LLPDCPAAAGCPPPPQAVGSGEVQNLTCGRNLSDFLVKTYPRLVROGLTKKWWNEVRY 1368
 Db 1473 MLPVCPGAGGLPPQPKQNTADILQDLTGRNLSDFLVKTYPRLVROGLTKKWWNEVRY 1532
 Qy 1369 GGFSLG--GRDPLPSQGLGRSVEELWALLSPLPGGALDRVLKMLTAWAHSLSDAQDSLKI 1427

Db	1533	GGFSLGVSNTQALPPSQSEVNDAIKQMKKHLKLAKDSADRFLNSLGRFTMTGLTRNNKV	1592
Qy	1428	WNNKNGHSMVAFVNRASNAILRAHLPPGPABHAHSITTLNHPNLNLTKEOLSEAAWLASS	1487
Db	1593	WNNKGWHAJSSPUNJUNNALLRANLQKGNPSHYGITAENHPNLNLTQKQOLSEVALMUTTS	1652
Qy	1488	VDLVLSICVVFAMSFVPASFTVLIEBERVTRAKHLQLMGGLSPTLYWLNFLMDMCNLYL	1547
Db	1653	VDLVLSICVIFAMSFVPASFWVFLOJQERVSXKXHLQFISGVKPYIWLNSFNWDMCNVYV	1712
Qy	1548	PACIVWLIFIAQORAYVAPANIPALLLLLLLLXYGWSITPLMWYPASFFSPSTAYVYLTC	1607
Db	1713	PATLVIIIFICFOQKSVYSNTLPVLALLLLLLXYGWSITPLMWYPASFEKIPSTAYVVLTS	1772
Qy	1608	INLPIGINGSMAFTVLELPSDQKLOEVSRIILKQVFLIFPHFCLGRGLIDMVRNQAMADAF	1667
Db	1773	VNLFIGINGSVATVLELFTDNKLNNLNDILKSVFLIFPHFCLGRGLIDMVKNQAMADAL	1832
Qy	1668	ERLGDROFQSPLRWVGVKNIILAMVIOGPIFLFTLLLOHRSOLLPOQVRSLPLLGEED	1727
Db	1833	ERFGENRFVPLSPMDLVGRNLFAMAVEGVVFFLITVLIOQRFIRPRPVNAKLSPUNDED	1892
Qy	1728	EDVARERVRVQATQGDVLVLRNLTKVYRGORMPAVDRLCLGTPPCECFGLLGVNGAGK	1787
Db	1893	EDVREERQRIIDGGQNDIILEIKELTKIYRKRPKPAVDRI CVGIPPGCECFGLLGVNGAGK	1952
Qy	1788	TSTFRMVTGDTLASRGAVLAGHSVAREPSAAHLSMGYCQSDAIPELLTGTREHLELLAR	1847
Db	1953	SSTFKMLTGDTTVTRGDAPFNKNSILSNIHEVHQNGMGCQFQDAITELLTGTREHVEFFAL	2012
Qy	1848	LRGYPEAQVATAGSGUARIGLSWYADRPAGTYSGGNKRKLATALAVGDPVAVFLDEPT	1907
Db	2013	LRGYPEKEVGKVGEWARKLGLVKYGEKYGAGTYSGGNKRKLSTAMALIGPPPVPVFLDEPT	2072
Qy	1908	TGMDPSARRFLNWSLLAVVREGRSVMLTSHSMECEALCSRLAIWNGRPRCIGSPQHLLK	1967
Db	2073	TGMDPKARRFLNWCALSVMKEGRSVMLTSHSMECEALCTRMAIWMNGRPRCIGSVQHLLK	2132
Qy	1968	GRFAAGHTLRLVPAARS--QPAAAFVAAEFPGSELREAHGGRRLFOLPGGRCALARVF	2025
Db	2133	NRFGDGYTVVRTAGSNPDILKPQDFFLGAFPGSVLKEKRNMLQQL--PSSLUSSLIARIF	2191
Qy	2026	GELAVHGAEGHVEDFSYSQTMLEBEVFLYFSKQOKDE--DTEBQKEAGVGVDPAFGLQH	2082
Db	2192	SILSQSKRKLHIEDYSYSQTTLDQVFNFAKQSDDDHLKDLSLHKNQTV--VDVAV----	2246
Qy	2083	PKRVSQFLDDPSTAET 2098	
Db	2247	--LTSFLODEKVKES 2259	

Search completed: January 3, 2005, 10:53:39
Job time : 189.061 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:28:54 ; Search time 108.665 Seconds
(without alignments)
7104.204 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQMLLLKQFMFYRRR.....QHPKRVSFQLDPPSTAETVL 2146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	11143	100.0	2146	9	US-09-995-542-5
2	11130	99.9	2146	15	US-10-182-006-2
3	11130	99.9	2180	15	US-10-332-447-27
4	11121.5	99.8	2147	17	US-10-618-281-44
5	11107	99.7	2144	9	US-09-858-194-2
6	11107	99.7	2144	14	US-10-154-419-2
7	10896	97.8	2100	9	US-09-995-542-6
8	10639.5	95.5	2059	15	US-10-114-270-176
9	9460.5	84.9	1873	15	US-10-182-006-4
10	8613.5	77.3	2167	9	US-09-995-542-2
11	8319.5	75.3	2121	9	US-09-995-542-3
12	7596	68.2	1550	9	US-09-995-542-8
13	5768.5	51.8	2261	14	US-10-313-641-9

14	5768.5	51.8	2261	14	US-10-313-641-10	Sequence 10, Appl
15	5768.5	51.8	2261	14	US-10-428-551-9	Sequence 9, Appl
16	5768.5	51.8	2261	14	US-10-428-551-10	Sequence 10, Appl
17	5768.5	51.8	2261	15	US-10-452-510-1	Sequence 1, Appl
18	5768.5	51.8	2261	15	US-10-617-334-1	Sequence 1, Appl
19	5768.5	51.8	2261	16	US-10-745-377-5	Sequence 5, Appl
20	5768.5	51.8	2261	16	US-10-744-465-1	Sequence 1, Appl
21	5768.5	51.8	2261	16	US-10-833-679-1	Sequence 1, Appl
22	5768.5	51.8	2261	17	US-10-872-113-5	Sequence 5, Appl
23	5752.5	51.6	2261	10	US-09-984-827-128	Sequence 128, App
24	5752.5	51.6	2261	10	US-09-984-827-132	Sequence 132, App
25	5752.5	51.6	2261	10	US-09-984-827-134	Sequence 134, App
26	5752.5	51.6	2261	10	US-09-984-827-136	Sequence 136, App
27	5751.5	51.6	2261	9	US-09-995-542-11	Sequence 11, Appl
28	5751.5	51.6	2261	9	US-09-845-456-11	Sequence 11, Appl
29	5751.5	51.6	2261	10	US-09-984-827-2	Sequence 2, Appl
30	5751.5	51.6	2261	10	US-09-984-827-127	Sequence 127, App
31	5751.5	51.6	2261	10	US-09-984-827-129	Sequence 129, App
32	5751.5	51.6	2261	10	US-09-984-827-130	Sequence 130, App
33	5751.5	51.6	2261	10	US-09-984-827-131	Sequence 131, App
34	5751.5	51.6	2261	10	US-09-984-827-135	Sequence 135, App
35	5747.5	51.6	2261	14	US-10-340-097-118	Sequence 118, App
36	5747.5	51.6	2261	14	US-10-336-215-118	Sequence 118, App
37	5747.5	51.6	2261	14	US-10-336-219-118	Sequence 118, App
38	5746.5	51.6	2261	10	US-09-984-827-133	Sequence 133, App
39	5744.5	51.6	2263	15	US-10-276-774-2326	Sequence 2326, Ap
40	5531.5	49.6	2201	9	US-09-995-542-9	Sequence 9, Appl
41	5522.5	49.6	2201	14	US-10-170-385-293	Sequence 293, App
42	5522.5	49.6	2201	14	US-10-331-496A-29	Sequence 29, Appl
43	5522.5	49.6	2201	15	US-10-429-160-4	Sequence 4, Appl
44	5522.5	49.6	2201	16	US-10-465-498-139	Sequence 139, App
45	5470.5	49.1	2233	16	US-10-465-498-140	Sequence 140, App

ALIGNMENTS

RESULT 1
US-09-995-542-5
; Sequence 5, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-5

Query Match	100.0%	Score	11143;	DB	9;	Length	2146;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2146;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAFWTQMLLLKQFMFYRRRQPVQLLVLLPFLFFLVAVRHSHHPLEHHECHFPNKP	60				
Db	1	MAFWTQMLLLKQFMFYRRRQPVQLLVLLPFLFFLVAVRHSHHPLEHHECHFPNKP	60				
Qy	61	LPSAGTVPWLQGLICNVNNTCFPOLTPOGEEPRLSNFNDLSVRLLDARTVLGGASAH	120				
Db	61	LPSAGTVPWLQGLICNVNNTCFPOLTPOGEEPRLSNFNDLSVRLLDARTVLGGASAH	120				
Qy	121	TLAGLGLKIATRAARSTAPQOPTKQSPLEPPMLDVAELLTSLLRATESLGLALGOQBPL	180				
Db	121	TLAGLGLKIATRAARSTAPQOPTKQSPLEPPMLDVAELLTSLLRATESLGLALGOQBPL	180				

Db 121 TLAGLGLKIATLRAARSTAOPTKQSPLEBPPMLDVAELLTSLRTESLGLALQAOEPL 180
Qy 181 HSLLEAAEDLAQELLALRSVELRALLQRPRGTSGPLELISEALCSVRGPSSTVGPSSLNW 240
Db 181 HSLLEAAEDLAQELLALRSVELRALLQRPRGTSGPLELISEALCSVRGPSSTVGPSSLNW 240
Qy 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDHPRLRLLWRLKPLILGKLLFAPD 300
Db 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDHPRLRLLWRLKPLILGKLLFAPD 300
Qy 301 TPFRKLMQAVNRRTFEBELTLRLDVRVWEMLGPRIFTFMDNSSNAMLQRLQWODEGRR 360
Db 301 TPFRKLMQAVNRRTFEBELTLRLDVRVWEMLGPRIFTFMDNSSNAMLQRLQWODEGRR 360
Qy 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAAPSEAA 420
Db 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAAPSEAA 420
Qy 421 LVSRALQLLAHRFRWAGVFLGPEDSDPTHEPTDLPDGPCHVRIRKIRMDIDVVRTNKIR 480
Db 421 LVSRALQLLAHRFRWAGVFLGPEDSDPTHEPTDLPDGPCHVRIRKIRMDIDVVRTNKIR 480
Qy 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYPCVYDD 540
Db 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYPCVYDD 540
Qy 541 VFLRVLSRSPLFLTLTAWIYSVTLTVKAAVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
Db 541 VFLRVLSRSPLFLTLTAWIYSVTLTVKAAVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
Qy 601 GPFLLSAAALLVILKLDGILPYSHPGVVFILAAFAVATVQSFLLSAFSTRANLAAACG 660
Db 601 GPFLLSAAALLVILKLDGILPYSHPGVVFILAAFAVATVQSFLLSAFSTRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVAVNRDLRPAAGRVAASLLSPVAFPGFCESLALLEOEGEAQWNVG 720
Db 661 GLAYFSLYLPVLCVAVNRDLRPAAGRVAASLLSPVAFPGFCESLALLEOEGEAQWNVG 720
Qy 721 TRPTADVLSLAQVSGLLLDALYGLATWYLEAVCPQGYGIPPEWPNPFRRSYWCGRPPP 780
Db 721 TRPTADVLSLAQVSGLLLDALYGLATWYLEAVCPQGYGIPPEWPNPFRRSYWCGRPPP 780
Qy 781 KSPAPCPTLPDVKVLEAPPGILSPGVSRSLKRPFGSPQALRGILSLDFOGHITAFIL 840
Db 781 KSPAPCPTLPDVKVLEAPPGILSPGVSRSLKRPFGSPQALRGILSLDFOGHITAFIL 840
Qy 841 GHNGAGKTTTILSLGILFPPSGGSATILGHDVRSMAAIRPHLCVCQYNNVLFDMLTVD 900
Db 841 GHNGAGKTTTILSLGILFPPSGGSATILGHDVRSMAAIRPHLCVCQYNNVLFDMLTVD 900
Qy 901 HWMPYGRKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGNQKLSVAIAFVGGSQV 960
Db 901 HWMPYGRKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGNQKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRGIGWELLKYREGRTILSTHLLDEAELLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRGIGWELLKYREGRTILSTHLLDEAELLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGYVILTVKARLPLTTNEKADTMEGSDVTRQEKNGSQSRVGTTPOLLA 1080
Db 1021 SPLFLRRHLGSGYVILTVKARLPLTTNEKADTMEGSDVTRQEKNGSQSRVGTTPOLLA 1080
Qy 1081 LVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLPRDLTRIAELRLTYGISTDTSLE 1140
Db 1081 LVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLPRDLTRIAELRLTYGISTDTSLE 1140
Qy 1141 EIFLKVVEECAADTMDGSGGCHLCTGAGLDVTLRLKMPPORTALENGEPAGSABETD 1200
Db 1141 EIFLKVVEECAADTMDGSGGCHLCTGAGLDVTLRLKMPPORTALENGEPAGSABETD 1200
Qy 1201 QGSGPDVAVRGVQWALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1260
Db 1201 QGSGPDVAVRGVQWALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1260

Qy 1261 PPFCHYPALRLSPMTYGAQVSPFSEDAPEGDPGRARLLEALLQEBAGLEEPVQVSHSSHRFSA 1320
Db 1261 PPFCHYPALRLSPMTYGAQVSPFSEDAPEGDPGRARLLEALLQEBAGLEEPVQVSHSSHRFSA 1320
Qy 1321 PEYPAEYAKVLASGNTPSPSPACQCSQPGARLLPDCPAAAGGPPPPQAVTSGSGVWQ 1380
Db 1321 PEYPAEYAKVLASGNTPSPSPACQCSQPGARLLPDCPAAAGGPPPPQAVTSGSGVWQ 1380
Qy 1381 NLTGRLNLDPLVKTYPRLVRQGLTKKVNNEVYGGFSLGGRDPGLSPGQELGRSVEBELW 1440
Db 1381 NLTGRLNLDPLVKTYPRLVRQGLTKKVNNEVYGGFSLGGRDPGLSPGQELGRSVEBELW 1440
Qy 1441 ALLSPLPGGALDRVLKXNLTAHAHSLDAQDSLSKTIWNNKNGHSMVAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKXNLTAHAHSLDAQDSLSKTIWNNKNGHSMVAFVNRASNAILRAHLP 1500
Qy 1501 PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSCVVVFAMSFVPASFTLVLEIE 1560
Db 1501 PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSCVVVFAMSFVPASFTLVLEIE 1560
Qy 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNLYLPACITVVLIFLAFQORAYVAPANLPALL 1620
Db 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNLYLPACITVVLIFLAFQORAYVAPANLPALL 1620
Qy 1621 LLLLLYCWSTITPLMYPASFPFSPSTAYVYVLTICINLFIGINGSMATFVLFSQDKLQEV 1680
Db 1621 LLLLLYCWSTITPLMYPASFPFSPSTAYVYVLTICINLFIGINGSMATFVLFSQDKLQEV 1680
Qy 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWEVVGKLLAMVIO 1740
Db 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWEVVGKLLAMVIO 1740
Qy 1741 GPLFLFTLLQLHRSQLLPQPRVRSPLFLGEEDEVARERERVVQATQGDVLVLRNLTK 1800
Db 1741 GPLFLFTLLQLHRSQLLPQPRVRSPLFLGEEDEVARERERVVQATQGDVLVLRNLTK 1800
Qy 1801 VYRGQMPAVDRCLGIPGECFGLLGVNGAGKSTFRMTVGTDLASRGAVALAGHSVAR 1860
Db 1801 VYRGQMPAVDRCLGIPGECFGLLGVNGAGKSTFRMTVGTDLASRGAVALAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVATQAGSLARLGLSWYAD 1920
Db 1861 EPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEAQVATQAGSLARLGLSWYAD 1920
Qy 1921 RPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLMNSLLAVVREGSRVML 1980
Db 1921 RPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLMNSLLAVVREGSRVML 1980
Qy 1981 TSHSMECECALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAPVAA 2040
Db 1981 TSHSMECECALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAPVAA 2040
Qy 2041 EPPGSELREAHGRLRFLQPLPGGRCALARVFGELAVHGAHGVDEFSVSTMLEEFLYF 2100
Db 2041 EPPGSELREAHGRLRFLQPLPGGRCALARVFGELAVHGAHGVDEFSVSTMLEEFLYF 2100
Qy 2101 SKDQGDDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEITVL 2146
Db 2101 SKDQGDDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEITVL 2146

RESULT 2

US-10-182-006-2
; Sequence 2, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23

1 PRIOR APPLICATION NUMBER: PCT/US01/02191
2 PRIOR FILING DATE: 2001-01-23
3 PRIOR APPLICATION NUMBER: 60/177,889
4 PRIOR FILING DATE: 2000-01-24
5 PRIOR APPLICATION NUMBER: 60/215,405
6 PRIOR FILING DATE: 2000-06-30
7 NUMBER OF SEQ ID NOS: 18
8 SOFTWARE: PatentIn Ver. 2.1
9 SEQ ID NO 2
10 LENGTH: 2146
11 TYPE: PRT
12 ORGANISM: Homo sapiens
13 US-10-182-006-2

Query Match 99.9%; Score 11130; DB 15; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2144; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MAFWTQMLLLKWNFMYRRRQVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP	60
DB	1	MAFWTQMLLLKWNFMYRRRQVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP	60
QY	61	LPSAGTVFWLQGLICNVANTCPQPTPGCEPGRLSNFNDLSVRLADARTVLGGASAH	120
DB	61	LPSAGTVFWLQGLICNVANTCPQPTPGCEPGRLSNFNDLSVRLADARTVLGGASAH	120
QY	121	TLAAGKLIATLARAARSTAOPTKQSPLEPMLDVAELLTSLRTESLGLALGOAEP	180
DB	121	TLAAGKLIATLARAARSTAOPTKQSPLEPMLDVAELLTSLRTESLGLALGOAEP	180
QY	181	HSLLEAAEDLAQELLALRSVELRALLQRPRTSGPFLLEALCSVRGPSSTVGPSLNW	240
DB	181	HSLLEAAEDLAQELLALRSVELRALLQRPRTSGPFLLEALCSVRGPSSTVGPSLNW	240
QY	241	YEASDLMELVGEPESALPDSSLSPACSELIGALDSHPILRLRRLKPLILGKLLFAPD	300
DB	241	YEASDLMELVGEPESALPDSSLSPACSELIGALDSHPILRLRRLKPLILGKLLFAPD	300
QY	301	TPFTKRLMAQVNRTEETLLRDREVWEMLGPRITFTFNDSSNVAMLRLLQMDQGR	360
DB	301	TPFTKRLMAQVNRTEETLLRDREVWEMLGPRITFTFNDSSNVAMLRLLQMDQGR	360
QY	361	QPRGGRDHEALRSFLDPGSGGYSWQDAHDVGLHVGTLGRVTECLSLDKLEAPSEAA	420
DB	361	QPRGGRDHEALRSFLDPGSGGYSWQDAHDVGLHVGTLGRVTECLSLDKLEAPSEAA	420
QY	421	LVSRALQLLAEHRFWAGVVFGLPESDPTTEHTPDLPFGHVRIRKIRMDIDVYTRTKIR	480
DB	421	LVSRALQLLAEHRFWAGVVFGLPESDPTTEHTPDLPFGHVRIRKIRMDIDVYTRTKIR	480
QY	481	DRFWDGPAADPLTLRLYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD	540
DB	481	DRFWDGPAADPLTLRLYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD	540
QY	541	VFLRVLRSPLFLFLTAIYVTLTKAVREKETRLDRMTKRAMGLSRAVLMLGWLFLSCL	600
DB	541	VFLRVLRSPLFLFLTAIYVTLTKAVREKETRLDRMTKRAMGLSRAVLMLGWLFLSCL	600
QY	601	GPFLLSAALLVLVLKGLDILPVSHPGVVFLLAFAVATVTSQPLLSAFFSRANLAAACG	660
DB	601	GPFLLSAALLVLVLKGLDILPVSHPGVVFLLAFAVATVTSQPLLSAFFSRANLAAACG	660
QY	661	GLAYFSLYLYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGAGHNVG	720
DB	661	GLAYFSLYLYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGAGHNVG	720
QY	721	TRPTADVFLSAQVSGLLLDALYGLATWYLEAVCPGYGIPPEPNWFFRYSYWCGRPP	780
DB	721	TRPTADVFLSAQVSGLLLDALYGLATWYLEAVCPGYGIPPEPNWFFRYSYWCGRPP	780
QY	781	KSPAPCPTPLDPKVLVEAPGLSPGVSRSLEKRFPGSPQALRGLSLDFYQGHITAPL	840
DB	781	KSPAPCPTPLDPKVLVEAPGLSPGVSRSLEKRFPGSPQALRGLSLDFYQGHITAPL	840

QY	841	GHNGAGKTTTILSILSGLPFPSPGSAFIIIGHDVRSSMAAIRPHLGYCPOYNVLFDMLTVD	900
DB	841	GHNGAGKTTTILSILSGLPFPSPGSAFIIIGHDVRSSMAAIRPHLGYCPOYNVLFDMLTVD	900
QY	901	HWFYGRUKLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSCGMQRKLSVAIAFVGGSV	960
DB	901	HWFYGRUKLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSCGMQRKLSVAIAFVGGSV	960
QY	961	VILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEAELLGDRVAVVAGRLCCCG	1020
DB	961	VILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEAELLGDRVAVVAGRLCCCG	1020
QY	1021	SPFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKNKSGQSGVGTPOQLLA	1080
DB	1021	SPFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKNKSGQSGVGTPOQLLA	1080
QY	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRLAELRLTGYISDTSLE	1140
DB	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRLAELRLTGYISDTSLE	1140
QY	1141	EFLKVVEECAADTDMEDSGCGHLCCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD	1200
DB	1141	EFLKVVEECAADTDMEDSGCGHLCCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD	1200
QY	1201	QSGGPDVAVRQGWALTRQOLQALLKRFLLARRRGLFAQIVLPALFVGLALVFSLIV	1260
DB	1201	QSGGPDVAVRQGWALTRQOLQALLKRFLLARRRGLFAQIVLPALFVGLALVFSLIV	1260
QY	1261	PPFGHYPALRLSPMTYGAQVFFSDEADPGDPCGRARLLLEALLQEAGLEPPVQSHSRFSA	1320
DB	1261	PPFGHYPALRLSPMTYGAQVFFSDEADPGDPCGRARLLLEALLQEAGLEPPVQSHSRFSA	1320
QY	1321	PEVPAEAKVILASGNWTPESPSPACQSCOPCARLLPCDPAAGPPPPQAVTGSVEVQ	1380
DB	1321	PEVPAEAKVILASGNWTPESPSPACQSCOPCARLLPCDPAAGPPPPQAVTGSVEVQ	1380
QY	1381	NLTGRNLSDFLVKTYPRLVROGLTKKWNEVRYGFSGLGRDPCGLPSGQELGRSVEELW	1440
DB	1381	NLTGRNLSDFLVKTYPRLVROGLTKKWNEVRYGFSGLGRDPCGLPSGQELGRSVEELW	1440
QY	1441	ALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIWNKNGHSMVAFVNRASNAIIRAHLP	1500
DB	1441	ALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIWNKNGHSMVAFVNRASNAIIRAHLP	1500
QY	1501	PGPARHAHSITTLNHPNLTKYQELSEALMASSVDVLVSIQVVFAMSFPASFTLVLTIE	1560
DB	1501	PGPARHAHSITTLNHPNLTKYQELSEALMASSVDVLVSIQVVFAMSFPASFTLVLTIE	1560
QY	1561	RVTRAKHLQMLGGLSPTLYWLGNFMDMKNLVACIVVLIPLAQOQRAYVAPANLPALL	1620
DB	1561	RVTRAKHLQMLGGLSPTLYWLGNFMDMKNLVACIVVLIPLAQOQRAYVAPANLPALL	1620
QY	1621	LLLLLYGWSITPLMYPASFPFSPSTAYVLTINLFIGINGSMTATVLELFSQKQJQEV	1680
DB	1621	LLLLLYGWSITPLMYPASFPFSPSTAYVLTINLFIGINGSMTATVLELFSQKQJQEV	1680
QY	1681	SRIILKQVFLIIPPHCLGRGLIDMVRNQAMADAFERLGRDQFOSPLRNEWGKNLLAMVIQ	1740
DB	1681	SRIILKQVFLIIPPHCLGRGLIDMVRNQAMADAFERLGRDQFOSPLRNEWGKNLLAMVIQ	1740
QY	1741	GPLFLLFTLLQHRSQLLPQPRVRSPLPGLGEDEVDARERVRVOCATQGDVLVLRNLT	1800
DB	1741	GPLFLLFTLLQHRSQLLPQPRVRSPLPGLGEDEVDARERVRVOCATQGDVLVLRNLT	1800
QY	1801	VYRQMPAVDRCLCLGIPPGCEFCGLLVNGAGKSTSTFRMVTGDTLASGEAVLAGHSVAR	1860
DB	1801	VYRQMPAVDRCLCLGIPPGCEFCGLLVNGAGKSTSTFRMVTGDTLASGEAVLAGHSVAR	1860
QY	1861	BPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWAD	1920
DB	1861	BPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWAD	1920


```
QY 1261 PFFGHYPALRLSPTMYGAOVSFSEDA PGDPGRARLEALLQEALEPPVQHSRRFSA 1320
Db 1295 PFFGHYPALRLSPTMYGAOVSFSEDA PGDPGRARLEALLQEALEPPVQHSRRFSA 1354
QY 1321 PEVPAEVAKVLASGNWTPSPSPACQCSOPGARLLLPDCPAAAGPPPPQAVTGSGEVVQ 1380
Db 1355 PEVPAEVAKVLASGNWTPSPSPACQCSOPGARLLLPDCPAAAGPPPPQAVTGSGEVVQ 1414
QY 1381 NLTGRNLSDFLVKTYPRVROGLKTKWNEVRYGFSI GGRDPGLPSGOELGRSVEELW 1440
Db 1415 NOTGRNLSDFLVKTYPRVROGLKTKWNEVRYGFSI GGRDPGLPSGOELGRSVEELW 1474
QY 1441 ALLSPLPGGALDRVLKNTLAWAHSDDAQDSLKIWFNNKGWHSWAFVNRASNAILRAHLP 1500
Db 1475 ALLSPLPGGALDRVLKNTLAWAHSDDAQDSLKIWFNNKGWHSWAFVNRASNAILRAHLP 1534
QY 1501 PGPARHAHSITITLNIHPLNLTKEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVLEE 1560
Db 1535 PGPARHAHSITITLNIHPLNLTKEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVLEE 1594
QY 1561 RVTRAKHLQLMGSLFTLVNLGNELWDMCNVLPACIVVLIIFLAFQORAYVAPANLPALL 1620
Db 1595 RVTRAKHLQLMGSLFTLVNLGNELWDMCNVLPACIVVLIIFLAFQORAYVAPANLPALL 1654
QY 1621 LLLLLYGWSITPLMTYPASGFFSVSPSTAYVVLTCINLFIGINGSMATFVLELPSDOKLOEV 1680
Db 1655 LLLLLYGWSITPLMTYPASGFFSVSPSTAYVVLTCINLFIGINGSMATFVLELPSDOKLOEV 1714
QY 1681 SRILKQVFLIPHFCLGRGLIDMVRNQAMADAFERLGRDQFQSPURWEVVGKLLAMVIQ 1740
Db 1715 SRILKQVFLIPHFCLGRGLIDMVRNQAMADAFERLGRDQFQSPURWEVVGKLLAMVIQ 1774
QY 1741 GFLFLLFTLLQHRSQLLPQPRVRSIPLLGEDEDAVARERERVVOGATQGDVILRNLTK 1800
Db 1775 GFLFLLFTLLQHRSQLLPQPRVRSIPLLGEDEDAVARERERVVOGATQGDVILRNLTK 1834
QY 1801 VYRGQMPAVDRCLGIPGECFGLLVNGAGKTSTFRMVTGDTLASGEAVLAGHSVAR 1860
Db 1835 VYRGQMPAVDRCLGIPGECFGLLVNGAGKTSTFRMVTGDTLASGEAVLAGHSVAR 1894
QY 1861 EPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLSWYAD 1920
Db 1895 EPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLSWYAD 1954
QY 1921 RPAGTYSGNKREKATALALVGDPAVVFLEDEPTTCMDPSARRFLWNSLLAVVREGSVM 1980
Db 1955 RPAGTYSGNKREKATALALVGDPAVVFLEDEPTTCMDPSARRFLWNSLLAVVREGSVM 2014
QY 1981 TSHSMECEALCSRLAIWVNGRFRCLGSPQHLKGRFAAGHTITLRVPAARSOPAAAFVAA 2040
Db 2015 TSHSMECEALCSRLAIWVNGRFRCLGSPQHLKGRFAAGHTITLRVPAARSOPAAAFVAA 2074
QY 2041 BPPGSELREAGGRRLRFPOLPGGRCALARVFGELAVHGAEGVEDFVSQTMLEVFLYF 2100
Db 2075 BPPGSELREAGGRRLRFPOLPGGRCALARVFGELAVHGAEGVEDFVSQTMLEVFLYF 2134
QY 2101 SKQCKDDETEQKEAGVGDPAVPCLOHPKRVSOFLDDPSTAEVT 2146
Db 2135 SKQCKDDETEQKEAGVGDPAVPCLOHPKRVSOFLDDPSTAEVT 2180
```

RESULT 4

US-10-618-281-44

; Sequence 44, Application US/10618281

; Publication No. US20040219609A1

; GENERAL INFORMATION:

; APPLICANT: Day, Anthony G.

; APPLICANT: Estell, David A.

; APPLICANT: Lyons, Eric H.

; APPLICANT: Yao, Jian

; TITLE OF INVENTION: Methods for Modulating Proteins Not

; PREVIOUSLY KNOWN AS: Proteases

```
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2147
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-618-281-44
```

```
Query Match 99.8%; Score 11121.5; DB 17; Length 2147;
Best Local Similarity 99.8%; Pred No. 0;
Matches 2143; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MAFWTQMLLWKQNFMYRRRQPVQLLVELLWPLFFILVAVRSHHPLEHHECHFPNKP 60
Db 1 MAFWTQMLLWKQNFMYRRRQPVQLLVELLWPLFFILVAVRSHHPLEHHECHFPNKP 60

QY 61 LPSAGTVPWLOGLICNVNNTCFPQPTGCEEPGRLSNFNDSLVSRLADARTVLGGASAH 120
Db 61 LPSAGTVPWLOGLICNVNNTCFPQPTGCEEPGRLSNFNDSLVSRLADARTVLGGASAH 120

QY 121 TLAGLGKLIATLRAARSTAOQPTKOSPPEPMLDVAELLTSLLRTESLGLALGOQBPL 180
Db 121 TLAGLGKLIATLRAARSTAOQPTKOSPPEPMLDVAELLTSLLRTESLGLALGOQBPL 180

QY 181 HSLLEAEADLAQELLALRSVLVELRALLQRPRTGSPLELSEALCSVRGSPSTVGPSLNW 240
Db 181 HSLLEAEADLAQELLALRSVLVELRALLQRPRTGSPLELSEALCSVRGSPSTVGPSLNW 240

QY 241 YEASDLMBELVQEPESALPDSLSLSPACSELGALDSHPSLRLRRLKPLILGKLLFAPD 300
Db 241 YEASDLMBELVQEPESALPDSLSLSPACSELGALDSHPSLRLRRLKPLILGKLLFAPD 300

QY 301 TPFTKRLMAQVNRIFEELTLDRVREWEMLGPRIFTEPMNDSSNVAMLQRLLOMDEGRR 360
Db 301 TPFTKRLMAQVNRIFEELTLDRVREWEMLGPRIFTEPMNDSSNVAMLQRLLOMDEGRR 360

QY 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGLGRVTECLSLDKLEAAPSEAA 420
Db 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGLGRVTECLSLDKLEAAPSEAA 420

QY 421 LVSRALQLLAHRFWAGVFLGPDSSDPTTCHPTDGLGPHVRKIRMDIDVTRTNKIR 480
Db 421 LVSRALQLLAHRFWAGVFLGPDSSDPTTCHPTDGLGPHVRKIRMDIDVTRTNKIR 480

QY 481 DRFWDPGPAADPLTDLRVWGGFVYVLODLVERAAVRVLSGANPRAGLYLOQMPYCYVDD 540
Db 481 DRFWDPGPAADPLTDLRVWGGFVYVLODLVERAAVRVLSGANPRAGLYLOQMPYCYVDD 540

QY 541 VFLRVLSRSLPLFTTLAWIYSVTITLVKAVVREKETRLDRMTWRAMGLSRAVLWGLWFLSCL 600
Db 541 VFLRVLSRSLPLFTTLAWIYSVTITLVKAVVREKETRLDRMTWRAMGLSRAVLWGLWFLSCL 600

QY 601 GPFLLSALLVLVLKLGDIILPSHPGVVFLAFAFATVTTQSFLLSAFFSRANLAAACG 660
Db 601 GPFLLSALLVLVLKLGDIILPSHPGVVFLAFAFATVTTQSFLLSAFFSRANLAAACG 660

QY 661 GLAYFSLYLPVLCVAMRDRLPAGRVAASLLSPVAFGFCESLALLEEQEGAGQHNNG 720
Db 661 GLAYFSLYLPVLCVAMRDRLPAGRVAASLLSPVAFGFCESLALLEEQEGAGQHNNG 720

QY 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWNPFFRKS YKCGRPP 780
Db 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWNPFFRKS YKCGRPP 780

QY 781 KSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRGSLDPIYQGHITAF 840
Db 781 KSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRGSLDPIYQGHITAF 840
```

```
Qy 841 CHNAGKTTTILSILSGLPPSGGSAFTLCHDVRSMAAIRPHLCVCPQYVNLFDMLTVDE 900
Db 841 GHNAGKTTTILSILSGLPPSGGSAFTLGHVDRSSMAAIRPHLCVCPQYVNLFDMLTVDE 900
Qy 901 HWFYGRLLKGLSAVAVGPEODRLQDVLVSQSVQTRHLSGGNORKLSVAIAFVGGSQV 960
Db 901 HWFYGRLLKGLSAVAVGPEODRLQDVLVSQSVQTRHLSGGNORKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDDEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRHLCGYVLTIVKARLPTTNEKADTMEGSDVTRQKKNGSQSRVGTPOLLA 1080
Db 1021 SPLFLRHLCGYVLTIVKARLPTTNEKADTMEGSDVTRQKKNGSQSRVGTPOLLA 1080
Qy 1081 LVQHWVPGARLVEELPHELVLVLPYTGADGSAFATLFDRLDTRLAELRLTGYSIDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLPYTGADGSAFATLFDRLDTRLAELRLTGYSIDTSLE 1140
Qy 1141 EIFLKVVVECAADTDMEDGSCQHLCTGIAGLDVTLRKMPPTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVVECAADTDMEDGSCQHLCTGIAGLDVTLRKMPPTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRSRGLFAQIVLPALFVGLAVFSLIV 1260
Db 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRSRGLFAQIVLPALFVGLAVFSLIV 1260
Qy 1261 PPFQHYPALRLSPMYGAQVVSFSEDAPGDPGRARLLEALLQEAAGLEPPVQVSHSRFSA 1320
Db 1261 PPFQHYPALRLSPMYGAQVVSFSEDAPGDPGRARLLEALLQEAAGLEPPVQVSHSRFSA 1320
Qy 1321 PEVPAEVAKLASGNWTPESPSPACQSCQPARLLPDCPAAAGPPPPQAVTSGSVEVQ 1380
Db 1321 PEVPAEVAKLASGNWTPESPSPACQSCQPARLLPDCPAAAGPPPPQAVTSGSVEVQ 1380
Qy 1381 NLTCGNLSDFLVKTYPRLVROGLTKKWVNEVRVCGFSLGGRDPLSGQELGRSVELW 1440
Db 1381 NLTCGNLSDFLVKTYPRLVROGLTKKWVNEVRVCGFSLGGRDPLSGQELGRSVELW 1440
Qy 1441 ALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGWSMVAFVNRAASNALTRAHL 1500
Db 1441 ALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGWSMVAFVNRAASNALTRAHL 1500
Qy 1501 PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSIQVFPAMSFVPASFTLVLIBE 1560
Db 1501 PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSIQVFPAMSFVPASFTLVLIBE 1560
Qy 1561 RVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPACIVVLIIFLAFQORAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPACIVVLIIFLAFQORAYVAPANLPALL 1620
Qy 1621 LLLLLYGWSITPLMPASFPFSPSTAYVLTICINLFIGINGSMAFVLELFSQ - QKLOE 1679
Db 1621 LLLLLYGWSITPLMPASFPFSPSTAYVLTICINLFIGINGSMAFVLELFSQ - QKLOE 1680
Qy 1680 VSRILKQVLIFFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKNLLAMVI 1739
Db 1681 VSRILKQVLIFFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKNLLAMVI 1740
Qy 1740 QGPLFLFTLLQHRSLIPQPRVSLPLIGEEDEDVAREERVVQCATQGDVLAFLNLT 1799
Db 1741 QGPLFLFTLLQHRSLIPQPRVSLPLIGEEDEDVAREERVVQCATQGDVLAFLNLT 1800
Qy 1800 KVRGQRMPAVDRLCLGIPGECFGLLVGNAGAKTSTFRMWTGDTLASRGEAVLAGHVA 1859
Db 1801 KVRGQRMPAVDRLCLGIPGECFGLLVGNAGAKTSTFRMWTGDTLASRGEAVLAGHVA 1860
Qy 1860 REPSAAHLSMGYCPQSDAIFELLTGRHELELLARLGRVPEAQVAQTAGSGLARLGLSWYA 1919
Db 1861 REPSAAHLSMGYCPQSDAIFELLTGRHELELLARLGRVPEAQVAQTAGSGLARLGLSWYA 1920
Qy 1920 DRPAGTSGGNKRLATLALVGPDPVAVFLDEPTTGMDPSARRFLWNSLLAVVREGSV 1979
```

```
Db 1921 DRPAGTSGGNKRLATLALVGPDPVAVFLDEPTTGMDPSARRFLWNSLLAVVREGSV 1980
Qy 1980 LTSHSMEECEALCSRLAIMVNGFRCLGSPQHUKGRFAAGHTTLTLRVPAARSQPAAFVA 2039
Db 1981 LTSHSMEECEALCSRLAIMVNGFRCLGSPQHUKGRFAAGHTTLTLRVPAARSQPAAFVA 2040
Qy 2040 AEPFGSELREAHGRLRFQPPGRCALARVFCGELAVHGAHEGVDESUSOTMLEEVFLY 2099
Db 2041 AEPFGSELREAHGRLRFQPPGRCALARVFCGELAVHGAHEGVDESUSOTMLEEVFLY 2100
Qy 2100 FSKDQKDEDETEQKAGVGVDPAPGLOHPKRVSQFLDDPSTAEITVL 2146
Db 2101 FSKDQKDEDETEQKAGVGVDPAPGLOHPKRVSQFLDDPSTAEITVL 2147

RESULT 5
US-09-858-194-2
; Sequence 2, Application US/09858194
; Patent No. US20020061590A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, MARIA
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF
; FILE REFERENCE: MNI-153
; CURRENT APPLICATION NUMBER: US/09/858,194
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/204,211
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-194-2

Query Match 99.7%; Score 11107; DB 9; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 MAFWTQMLLLWKNFMVRRRQPVQLLVLELWPLFLPFLVAVRSHSPPLHHECHFPNKP 60
Db 1 MAFWTQMLLLWKNFMVRRRQPVQLLVLELWPLFLPFLVAVRSHSPPLHHECHFPNKP 58
Qy 61 LPSAGTVPMLOGLICNVNNTCFPOLTPGEBPGRLSNFNDLSVSRLLADARTVLGGASAH 120
Db 59 LPSAGTVPMLOGLICNVNNTCFPOLTPGEBPGRLSNFNDLSVSRLLADARTVLGGASAH 118
Qy 121 TLAGLGLIATLRAARSTAOPTKOSPLEPPMLDVAELTSLRTESLGLALGQAOEPL 180
Db 119 TLAGLGLIATLRAARSTAOPTKOSPLEPPMLDVAELTSLRTESLGLALGQAOEPL 178
Qy 181 HSLEAAEDLAQELLALRSILVELRALLQRPRTSGPLELLELSEALCSVRGSPSTVGPSLNW 240
Db 179 HSLEAAEDLAQELLALRSILVELRALLQRPRTSGPLELLELSEALCSVRGSPSTVGPSLNW 238
Qy 241 YEASDLMELVGPESALPDSSISPACSELIGALDHSPLSRLLWRRLKPIILGKLLFAPD 300
Db 239 YEASDLMELVGPESALPDSSISPACSELIGALDHSPLSRLLWRRLKPIILGKLLFAPD 298
Qy 301 TPTRKLMQAVNTTFFELTLRLDRVREVMELGPRIETFMNDSSNVAMQLQLQODSGRR 360
Db 299 TPTRKLMQAVNTTFFELTLRLDRVREVMELGPRIETFMNDSSNVAMQLQLQODSGRR 358
Qy 361 QRPFGGRDHMEALRSFLDPCSGGYSWQDAHADVHVLVGLTGRVTECLSLDKLEAAPSEAA 420
Db 359 QRPFGGRDHMEALRSFLDPCSGGYSWQDAHADVHVLVGLTGRVTECLSLDKLEAAPSEAA 418
Qy 421 LVSRALQLLAHRAHFWAGVFLGPDSSDPTHEPTDPLGPGHVRIRKIRMDIVVTRTNKIR 480
Db 419 LVSRALQLLAHRAHFWAGVFLGPDSSDPTHEPTDPLGPGHVRIRKIRMDIVVTRTNKIR 478
```

QY 481 DRFDGPAADPLTDLRYVWGFFVYLQDLIVERAAVRVLSGANPRAGLYLQMPYCYVDD 540
DB 479 DRFDGPAADPLTDLRYVWGFFVYLQDLIVERAAVRVLSGANPRAGLYLQMPYCYVDD 538
QY 541 VFLRVLSRLPLFLTLAMIYVTLTVKAVVRKETRRLRDTMRAMGLSRAVLWLGWFLSCL 600
DB 539 VFLRVLSRLPLFLTLAMIYVTLTVKAVVRKETRRLRDTMRAMGLSRAVLWLGWFLSCL 598
QY 601 GPFLLSAALLVLVLDGILPYSHGCVVFLFLAAFAVATVVTQSFLLSAFFSRANLAAACG 660
DB 599 GPFLLSAALLVLVLDGILPYSHGCVVFLFLAAFAVATVVTQSFLLSAFFSRANLAAACG 658
QY 661 GLAYFSLYLYVLCVAMDRDLPAAGRVAASLLSPVAFGFCBSLALLBEEQGEAGWHNVG 720
DB 659 GLAYFSLYLYVLCVAMDRDLPAAGRVAASLLSPVAFGFCBSLALLBEEQGEAGWHNVG 718
QY 721 TRPTADVFSLAQVSGLLLDAAALYGLATWYLBVAVCPQGYGIPDPNPPFRSRYWCGPRPP 780
DB 719 TRPTADVFSLAQVSGLLLDAAALYGLATWYLBVAVCPQGYGIPDPNPPFRSRYWCGPRPP 778
QY 781 KSPAPCPTPLDPKVLVEBAPGCLSPGVSVRSLEKFPFGSPQALRGLSLDIFYQGHITAPL 840
DB 779 KSPAPCPTPLDPKVLVEBAPGCLSPGVSVRSLEKFPFGSPQALRGLSLDIFYQGHITAPL 838
QY 841 GHNGAGKTTTILSGLPFPSPGSAFIFLGHVDRSSMAAIRPHLGYCPOYNVLFDMLTVDY 900
DB 839 GHNGAGKTTTILSGLPFPSPGSAFIFLGHVDRSSMAAIRPHLGYCPOYNVLFDMLTVDY 898
QY 901 HWYFVGRKGLSAAVGPEDRLLDQVGLVSKQSVOTRHLSGMORKLSVAIAFVGGSQV 960
DB 899 HWYFVGRKGLSAAVGPEDRLLDQVGLVSKQSVOTRHLSGMORKLSVAIAFVGGSQV 958
QY 961 VTLDPTAGVDPASRRGIWELLKVKREGRTLLSTHLDDEALLGDRVAVVAGRLCCCG 1020
DB 959 VTLDPTAGVDPASRRGIWELLKVKREGRTLLSTHLDDEALLGDRVAVVAGRLCCCG 1018
QY 1021 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKNKGSGSRVGTPLL 1080
DB 1019 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKNKGSGSRVGTPLL 1078
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAELRLTGYISDTSLE 1140
DB 1079 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAELRLTGYISDTSLE 1138
QY 1141 EFLKVVBECAADTDMEGSCQHLCCTGIAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
DB 1139 EFLKVVBECAADTDMEGSCQHLCCTGIAGLDVTLRLKMPQETALENGEPAGSAPETD 1198
QY 1201 QGSGPDVGRVQGWALTRQQLALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1260
DB 1199 QGSGPDVGRVQGWALTRQQLALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1258
QY 1261 PPFHYPALRLSPTWYGAQVSPFSDADPGDGRARLLEALLOEAGLEPPVQVSHRPSA 1320
DB 1259 PPFHYPALRLSPTWYGAQVSPFSDADPGDGRARLLEALLOEAGLEPPVQVSHRPSA 1318
QY 1321 PEVPAEVAKLASGNWTPESPACQCPGARRLLPDCPAAAGCPPPPQAVTSGGEVVQ 1380
DB 1319 PEVPAEVAKLASGNWTPESPACQCPGARRLLPDCPAAAGCPPPPQAVTSGGEVVQ 1378
QY 1381 NLTRGNLSDFLVKTYPRLVROGLKTKKWVNEVRYGFSGLGRDPCGLPSQELGRSVEELW 1440
DB 1379 NLTRGNLSDFLVKTYPRLVROGLKTKKWVNEVRYGFSGLGRDPCGLPSQELGRSVEELW 1438
QY 1441 ALLSPLPGALDRVLKNTAWAHSDDAQSLSKIWFNNKGWSWAFVNRASNAIILRAHLP 1500
DB 1439 ALLSPLPGALDRVLKNTAWAHSDDAQSLSKIWFNNKGWSWAFVNRASNAIILRAHLP 1498
QY 1501 PGPARHAHSITTLNHLPLNLTKEQLSEALMASSVDVLVSI CVVFWMSFVPASFTLVLEE 1560
DB 1499 PGPARHAHSITTLNHLPLNLTKEQLSEALMASSVDVLVSI CVVFWMSFVPASFTLVLEE 1558
QY 1561 RVTRAKHLQMGGLSPTLYLWGNFDMCNVLPACI VVLIPLAQRAYVAPANLPALL 1620

DB 1559 RVTRAKHLQMGGLSPTLYLWGNFDMCNVLPACI VVLIPLAQRAYVAPANLPALL 1618
QY 1621 LLLLYLXGHSITPLMYPASFFSVSTAYVVLTCINLPIFGINGSMATFVLELFSQKLOQV 1680
DB 1619 LLLLYLXGHSITPLMYPASFFSVSTAYVVLTCINLPIFGINGSMATFVLELFSQKLOQV 1678
QY 1681 SRLKQVFLIIPPHFCLGRGLIDMVNRQAMADAFERLGRQFOSPLRWEVVGKNLLAMVIO 1740
DB 1679 SRLKQVFLIIPPHFCLGRGLIDMVNRQAMADAFERLGRQFOSPLRWEVVGKNLLAMVIO 1738
QY 1741 GPLFLLTLLLOHRSQLLPQPRVRSPLPLGBEDEDAVARERERVVQATQGDVVLRLNLT 1800
DB 1739 GPLFLLTLLLOHRSQLLPQPRVRSPLPLGBEDEDAVARERERVVQATQGDVVLRLNLT 1798
QY 1801 VYRGQMPAVDRCLCIGIPGECFGLLVNGAGKTSITFRMVTGDTLASGEAVLAGHSVAR 1860
DB 1799 VYRGQMPAVDRCLCIGIPGECFGLLVNGAGKTSITFRMVTGDTLASGEAVLAGHSVAR 1858
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVAGTAGSGLARLGLSWAD 1920
DB 1859 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVAGTAGSGLARLGLSWAD 1918
QY 1921 RPAGTYSGNKRKLATALALVGDPAVVFDEPTTGMDSARRFLWNSLLAVVREGSVML 1980
DB 1919 RPAGTYSGNKRKLATALALVGDPAVVFDEPTTGMDSARRFLWNSLLAVVREGSVML 1978
QY 1981 TSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSOPAAAFVAA 2040
DB 1979 TSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSOPAAAFVAA 2038
QY 2041 EFPQSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEBFLYF 2100
DB 2039 EFPQSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEBFLYF 2098
QY 2101 SKDQKQEDTEEQKEAGVGDVPAPGLQHPKRVSQFLDDPSTAEVTL 2146
DB 2099 SKDQKQEDTEEQKEAGVGDVPAPGLQHPKRVSQFLDDPSTAEVTL 2144

RESULT 6
US-10-154-419-2
; Sequence 2, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-2

Query Match 99.7%; Score 11107; DB 14; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
QY 1 MAFVTQLMLLLWKFMVRRRROPVOLLVLELLWPLFLFFTLVAVRSHSPPLEHHECHFPNKP 60
DB 1 MAFVTQLMLLLWKFMVRRRROPVOLLVLELLWPLFLFFTLVAVRSHSPPLEHHECHFPNKP 58
QY 61 LPSAGTVPLWQGLICNVNNTCFPQLTPGEEPRGLSNFNDLSVSLRADARTVLCGASAH 120

Db 59 LPSAGTVPWLOGLICNVNNTCFPOLITGEEBGRSLNFENDSLVSRLLADARTVLGASAH 118
Qy 121 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTESLGLAQOEPL 180
Db 119 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTESLGLAQOEPL 178
Qy 181 HSILEAAEDLAEILLALSLVELLALLORPGTSGPLELSEALCSVRGSPSTVGPSSLNW 240
Db 179 HSILEAAEDLAEILLALSLVELLALLORPGTSGPLELSEALCSVRGSPSTVGPSSLNW 238
Qy 241 YEASDLMELVQEPESALPDSSLSPACSELIGALDHPLSRLWRLKPLILGKLLFAPD 300
Db 239 YEASDLMELVQEPESALPDSSLSPACSELIGALDHPLSRLWRLKPLILGKLLFAPD 298
Qy 301 TPFTRKMAQVNRFTFEELTLARDVREWEMLGPRIFTFMDSSNVAMLQRLQWQDEGRR 360
Db 299 TPFTRKMAQVNRFTFEELTLARDVREWEMLGPRIFTFMDSSNVAMLQRLQWQDEGRR 358
Qy 361 QPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Db 359 QPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 418
Qy 421 LVSRALQLLAHRFWAGVVFGLPEDSSDPTHEPTDLPQGHVRIKIRMDIDVTRTKIR 480
Db 419 LVSRALQLLAHRFWAGVVFGLPEDSSDPTHEPTDLPQGHVRIKIRMDIDVTRTKIR 478
Qy 481 DRFWDGPAADPLTDLRYVMGGFYVLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
Db 479 DRFWDGPAADPLTDLRYVMGGFYVLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 538
Qy 541 VFLRVLSRSLPLFLTLAWISVTLTVKAVVREKETRLRDTWRANGLSRAVLWLGWFLSCL 600
Db 539 VFLRVLSRSLPLFLTLAWISVTLTVKAVVREKETRLRDTWRANGLSRAVLWLGWFLSCL 598
Qy 601 GPFLLSAALLVLKLGDIILPYSHPGVVFLLAFAFVATVTSFLLSAFFSRANLAAACG 660
Db 599 GPFLLSAALLVLKLGDIILPYSHPGVVFLLAFAFVATVTSFLLSAFFSRANLAAACG 658
Qy 661 GLAYFSLYLPLYVLCVAMRDLRDPAGGRVAASLLSPVAFGFCESLALLEQEGEQAHNVG 720
Db 659 GLAYFSLYLPLYVLCVAMRDLRDPAGGRVAASLLSPVAFGFCESLALLEQEGEQAHNVG 718
Qy 721 TRPTADVFLSAQVSGLLLLDAAALGLATWYLEAVCPQGYGIPPEWNPFFRRSYWCGRPP 780
Db 719 TRPTADVFLSAQVSGLLLLDAAALGLATWYLEAVCPQGYGIPPEWNPFFRRSYWCGRPP 778
Qy 781 KSPAPCPTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPOPALRGLSLDFYQGHITAF 840
Db 779 KSPAPCPTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPOPALRGLSLDFYQGHITAF 838
Qy 841 GHNGAGKTTTLLSILSGLFPFGSGSAFILGHHDVRSMAAIRPHLGCPCQYNVLFDMLTVD 900
Db 839 GHNGAGKTTTLLSILSGLFPFGSGSAFILGHHDVRSMAAIRPHLGCPCQYNVLFDMLTVD 898
Qy 901 HWFYGRLKGLSAAVGEQDRLLQDVLVSQSVQTRHLSGQNRKLSVAIAIFVGSSQV 960
Db 899 HWFYGRLKGLSAAVGEQDRLLQDVLVSQSVQTRHLSGQNRKLSVAIAIFVGSSQV 958
Qy 961 VILDEPTAGVDPASRRGIWELLKYRGRGRTLILSTHLLDEALLGDRVAVVAGGRGCCG 1020
Db 959 VILDEPTAGVDPASRRGIWELLKYRGRGRTLILSTHLLDEALLGDRVAVVAGGRGCCG 1018
Qy 1021 SPLFLRRLHLSGGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKNGSQSGSRVGTPOLLA 1080
Db 1019 SPLFLRRLHLSGGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKNGSQSGSRVGTPOLLA 1078
Qy 1081 LVQHWVPCARLVEELPHELVLVILPYTCANDGSFATLIFRELDTRLAEILRTCYGISTDSLE 1140
Db 1079 LVQHWVPCARLVEELPHELVLVILPYTCANDGSFATLIFRELDTRLAEILRTCYGISTDSLE 1138
Qy 1141 EIFLKVVEECAADTDMEDGSCQGLCTGIAGLDVTLRLKMPPOBTALENGEPAGSAPETD 1200

Db 1139 EIFLKVVEECAADTDMEDGSCQGLCTGIAGLDVTLRLKMPPOBTALENGEPAGSAPETD 1198
Qy 1201 QSGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGPAQIVLPALFVGLALVFSLI 1260
Db 1199 QSGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGPAQIVLPALFVGLALVFSLI 1258
Qy 1261 PPGHYPALRLSPMTYCAVUSFFSEADPGDPGRARLLLEALLQAGLEPEEPVQHSSHRFSA 1320
Db 1259 PPGHYPALRLSPMTYCAVUSFFSEADPGDPGRARLLLEALLQAGLEPEEPVQHSSHRFSA 1318
Qy 1321 PEYPAEVAKVLAGSNWTPSPSPACQSQPGARLLPDCPAAAGGPPPPQAVTSGSGVQ 1380
Db 1319 PEYPAEVAKVLAGSNWTPSPSPACQSQPGARLLPDCPAAAGGPPPPQAVTSGSGVQ 1378
Qy 1381 NLFGRLNLSDFLVKTYPRLVQRGLTKKVNVEVYGGFSGLGRDPGLPSGELGRSVEELW 1440
Db 1379 NLFGRLNLSDFLVKTYPRLVQRGLTKKVNVEVYGGFSGLGRDPGLPSGELGRSVEELW 1438
Qy 1441 ALLSPLPGGALDRVLKNTLAWHSLDAODSLKTFWNNKGWHSVAFVNRASNAIILRAHLP 1500
Db 1439 ALLSPLPGGALDRVLKNTLAWHSLDAODSLKTFWNNKGWHSVAFVNRASNAIILRAHLP 1498
Qy 1501 PGPARHAHSITTLNHPNLNTKEOLSEAAALMASSVDVLVSI CVVFAMSFVPASFTLVLEE 1560
Db 1499 PGPARHAHSITTLNHPNLNTKEOLSEAAALMASSVDVLVSI CVVFAMSFVPASFTLVLEE 1558
Qy 1561 RVTRAKHLQMGSLSPTLVWLGNFMDMCMNYLPACIVVLIFLAFOORATVAPANLPALL 1620
Db 1559 RVTRAKHLQMGSLSPTLVWLGNFMDMCMNYLPACIVVLIFLAFOORATVAPANLPALL 1618
Qy 1621 LLLLLXGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEV 1680
Db 1619 LLLLLXGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLOEV 1678
Qy 1681 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLQDQFQSPLRWEVVGKLLAMVIQ 1740
Db 1679 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLQDQFQSPLRWEVVGKLLAMVIQ 1738
Qy 1741 GPLFLFTLLLOHRSQLLPOPRVRSPLPGEDEEDVARERERVVQATQGDVLRNLTK 1800
Db 1739 GPLFLFTLLLOHRSQLLPOPRVRSPLPGEDEEDVARERERVVQATQGDVLRNLTK 1798
Qy 1801 VYRGORMPADVRLCLGIPGCEGFLGNGAGKSTFRMVGTDLASRGEAVLAGHSVAR 1860
Db 1799 VYRGORMPADVRLCLGIPGCEGFLGNGAGKSTFRMVGTDLASRGEAVLAGHSVAR 1858
Qy 1861 EPSAAHLSMGYCQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1920
Db 1859 EPSAAHLSMGYCQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1918
Qy 1921 RPAGTYSGGNKRKLATALLVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSRVML 1980
Db 1919 RPAGTYSGGNKRKLATALLVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSRVML 1978
Qy 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPHOLKAGRFAAGHTLTLRVPAARSQAAAAVAA 2040
Db 1979 TSHSMECEALCSRLAIMVNGRFRCLGSPHOLKAGRFAAGHTLTLRVPAARSQAAAAVAA 2038
Qy 2041 EPPGSELREAHGRLRFLQPLPGGRCALARVFGELAVHGAHEGVDEFSVQTMLEEFLYF 2100
Db 2039 EPPGSELREAHGRLRFLQPLPGGRCALARVFGELAVHGAHEGVDEFSVQTMLEEFLYF 2098
Qy 2101 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEITVL 2146
Db 2099 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEITVL 2144

RESULT 7
US-09-995-542-6
; Sequence 6, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John

APPLICANT: Ulias, Jaarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR FILING DATE: 2001-11-28
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-6

Query Match 97.8%; Score 10896; DB 9; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	47	PPEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCPQOLTPGEEGRSLNFNDSLSVRL	106
DB	1	PPEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCPQOLTPGEEGRSLNFNDSLSVRL	60
QY	107	ADARTVLGGASAHRTLAGLGLIATLRAARSTAQPOPTKQSPLEPPMLDVAELLTSLRT	166
DB	61	ADARTVLGGASAHRTLAGLGLIATLRAARSTAQPOPTKQSPLEPPMLDVAELLTSLRT	120
QY	167	ESIGLALGOAQPPLSHLLBAEDLAQELLALSLVELRALLORPGTSGPLELLSEALCS	226
DB	121	ESIGLALGOAQPPLSHLLBAEDLAQELLALSLVELRALLORPGTSGPLELLSEALCS	180
QY	227	VRGSPSTVGPSLNWYEASDLMEIVGQEPESALPDSSLSPACSELICGALDSDHPLSLMWR	286
DB	181	VRGSPSTVGPSLNWYEASDLMEIVGQEPESALPDSSLSPACSELICGALDSDHPLSLMWR	240
QY	287	LKPLILGKLLFAPDPTFKLMAQVNRITPEELTLRDVREVWMLGPRIFTFMDSSNVA	346
DB	241	LKPLILGKLLFAPDPTFKLMAQVNRITPEELTLRDVREVWMLGPRIFTFMDSSNVA	300
QY	347	MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQADHADVGLVGLGRVTEC	406
DB	301	MLQRLQWQDEGRQPRPGGRDHMEALRSFLDPGSGGYSWQADHADVGLVGLGRVTEC	360
QY	407	LSLDKLEAPSAALVSRAQLAHRFWAGVVLGPDSSDPTHEPTDLPDGPVRIKI	466
DB	361	LSLDKLEAPSAALVSRAQLAHRFWAGVVLGPDSSDPTHEPTDLPDGPVRIKI	420
QY	467	RMDIDVTRTNKIRDFWDPGPAADPLTDLRYVMGGFVYLDLVERAAVRVLSGANPRAG	526
DB	421	RMDIDVTRTNKIRDFWDPGPAADPLTDLRYVMGGFVYLDLVERAAVRVLSGANPRAG	480
QY	527	LYLQMPYPCYVDDVFLVLSRLSPLFTLAMIYSVTITVKAIVREKETRLDTRWAMGL	586
DB	481	LYLQMPYPCYVDDVFLVLSRLSPLFTLAMIYSVTITVKAIVREKETRLDTRWAMGL	540
QY	587	SPAVLWGLFELCLGPFLLSALALVILKGLDILPSHPGVFLPLAFAVATVTSFLL	646
DB	541	SPAVLWGLFELCLGPFLLSALALVILKGLDILPSHPGVFLPLAFAVATVTSFLL	600
QY	647	SAFFSRANLAACGGGLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLAL	706
DB	601	SAFFSRANLAACGGGLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLAL	660
QY	707	LEEQEGQAQHNVTGRTADVPSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWN	766
DB	661	LEEQEGQAQHNVTGRTADVPSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWN	720
QY	767	PPFRSYNGRRPPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPQPALRG	826
DB	721	PPFRSYNGRRPPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRFPGSPQPALRG	780
QY	827	LSLDFYQGHITAFLGHNGAGKTTTLLSGLFPPSGGSFAFLGHHDVRSMAAIRPHLVC	886

DB	781	LSLDFYQGHITAFLGHNGAGKTTTLLSGLFPPSGGSFAFLGHHDVRSMAAIRPHLVC	840
QY	887	PQYNVLFDMLTVDHWFYGRUKGLSAAVGPEDORLLQDVLGVSKQSVOTRHLSCGMQR	946
DB	841	PQYNVLFDMLTVDHWFYGRUKGLSAAVGPEDORLLQDVLGVSKQSVOTRHLSCGMQR	900
QY	947	KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREGRTILSLTHHLEAEELGD	1006
DB	901	KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREGRTILSLTHHLEAEELGD	960
QY	1007	RVAVVAGRLCCCGSPLFLRRHLGSGYILTLVKARLPLTTNEKADTDMEGSVDTREKKN	1066
DB	961	RVAVVAGRLCCCGSPLFLRRHLGSGYILTLVKARLPLTTNEKADTDMEGSVDTREKKN	1020
QY	1067	SGQSRVGTPOLLALVQHWVPGARLVEBELPHELVLVLVPTGAHDSGFATLPRELDTLAE	1126
DB	1021	SGQSRVGTPOLLALVQHWVPGARLVEBELPHELVLVLVPTGAHDSGFATLPRELDTLAE	1080
QY	1127	LRLTCYGISDTSLEEIFLKVVVEECAADTDMEGSCGHLCTGCIAGLDVTLRLKMPQETA	1186
DB	1081	LRLTCYGISDTSLEEIFLKVVVEECAADTDMEGSCGHLCTGCIAGLDVTLRLKMPQETA	1140
QY	1187	LENGEPAGSAPETDQSGPDVAVRQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP	1246
DB	1141	LENGEPAGSAPETDQSGPDVAVRQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP	1200
QY	1247	ALFVGLALVFSILVPPFGHYPALRLSPMTYGAQVFFSDEADPGDPCGRARLLLEALQEAGL	1306
DB	1201	ALFVGLALVFSILVPPFGHYPALRLSPMTYGAQVFFSDEADPGDPCGRARLLLEALQEAGL	1260
QY	1307	EEPPVQSHSHRFAPEVAEVAKVLASGNWTPESSPACQSCQSGARLLPDCPAAAGP	1366
DB	1261	EEPPVQSHSHRFAPEVAEVAKVLASGNWTPESSPACQSCQSGARLLPDCPAAAGP	1320
QY	1367	PPQAVTSGGEVQMLTGRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGFSLGRDPL	1426
DB	1321	PPQAVTSGGEVQMLTGRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGFSLGRDPL	1380
QY	1427	PSGQELGRSVBELWALLSPLPGGALDRVLKNTAWAHLDAQDSLKIWFNNKGHSMWAF	1486
DB	1381	PSGQELGRSVBELWALLSPLPGGALDRVLKNTAWAHLDAQDSLKIWFNNKGHSMWAF	1440
QY	1487	VNRASNALRAHLPPGPARHAHSITTLNHPNLNLTKEQSEAALMASSVDVLVSIQVVFAM	1546
DB	1441	VNRASNALRAHLPPGPARHAHSITTLNHPNLNLTKEQSEAALMASSVDVLVSIQVVFAM	1500
QY	1547	SFVPASFTLVLIERVTAKHLQMLGGLSPTLYWLGNFWDNCNVLVPACIVVLIFLAFQ	1606
DB	1501	SFVPASFTLVLIERVTAKHLQMLGGLSPTLYWLGNFWDNCNVLVPACIVVLIFLAFQ	1560
QY	1607	QRAYVAPANLALLLLLYGWSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMAT	1666
DB	1561	QRAYVAPANLALLLLLYGWSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMAT	1620
QY	1667	FVLELFSQKLOEVSRIILKQVFLIFPHECLGRGLIDMVRNQAMADAFERLGRQSPPLR	1726
DB	1621	FVLELFSQKLOEVSRIILKQVFLIFPHECLGRGLIDMVRNQAMADAFERLGRQSPPLR	1680
QY	1727	WEVVGKLLAMVIOGPIFLFTLLLOHRSQLLPOPRVRSPLLGEEDDVARERERVQ	1786
DB	1681	WEVVGKLLAMVIOGPIFLFTLLLOHRSQLLPOPRVRSPLLGEEDDVARERERVQ	1740
QY	1787	ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGCEGGLGVNGAGKTTSTFRMTVGTDLA	1846
DB	1741	ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPGCEGGLGVNGAGKTTSTFRMTVGTDLA	1800
QY	1847	SRGEAVLAGHSVAREPSNAHLSMGCQSDAIFELLTGREHLEILLARLGRVPEAQVATA	1906
DB	1801	SRGEAVLAGHSVAREPSNAHLSMGCQSDAIFELLTGREHLEILLARLGRVPEAQVATA	1860
QY	1907	GSGLARLGLSNYADRPAGTYSGGNKRKLATALVGDPAVVFLEPTTGMDBPSARRFLWN	1966

Db 1861 GSGIARLGLSWYADRPAGTSGGNKRKLATALALVGDPAVFLDEPTTGMDPSARRFLWN 1920
Qy 1967 SLLAVVREGSRVMTLSHMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 2026
Db 1921 SLLAVVREGSRVMTLSHMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Qy 2027 PAARSQAPAAVFAAEFFGSELRREAHGGRRLRFPQLPPGRCALARVFGELAVHGAHGVEDF 2086
Db 1981 PAARSQAPAAVFAAEFFGSELRREAHGGRRLRFPQLPPGRCALARVFGELAVHGAHGVEDF 2040
Qy 2087 SVSOTMLEEVFLYFSKQGDDETEEOKEAGVGVDPAPGLQHPKRVSQFLDDPSTASTVL 2146
Db 2041 SVSOTMLEEVFLYFSKQGDDETEEOKEAGVGVDPAPGLQHPKRVSQFLDDPSTASTVL 2100

RESULT 8

US-10-114-270-176
; Sequence 176, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaozhong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Murallidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234

; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 176
; LENGTH: 2059
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-114-270-176

Query Match 95.5%; Score 10639.5; DB 15; Length 2059;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2058; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 MAFWTQLMLLWKNFMYRRRQPVQLLVLLWLPFLFFILVAVRHSHPPLHHECHFPNKP 60
Db 1 MAFWTQLMLLWKNFMYRRRQPVQLLVLLWLPFLFFILVAVRHSHPPLHHECHFPNKP 60
Qy 61 LPSAGTVPMLOGLICNVNNTCFPQLTPGPEPGRLSNPNDSLVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPMLOGLICNVNNTCFPQLTPGPEPGRLSNPNDSLVSRLLADARTVLGGASAH 120
Qy 121 TLAGLGKLIATLRAARSTAOPTKOSPPLPPMDVAELTSLRTESLGLAQOEPL 180
Db 121 TLAGLGKLIATLRAARSTAOPTKOSPPLPPMDVAELTSLRTESLGLAQOEPL 180
Qy 181 HSLEAAEDIAQELLALRSIVELRALLORPRGTSGPPELLSEALCSVRGSSSTVGPSSLN 240
Db 181 HSLEAAEDIAQELLALRSIVELRALLORPRGTSGPPELLSEALCSVRGSSSTVGPSSLN 240
Qy 241 YEASDLMEVGQEPESALPDSSISPACSELIGALDSHPISRLLRRLKPIILGKLLFAPD 300
Db 241 YEASDLMEVGQEPESALPDSSISPACSELIGALDSHPISRLLRRLKPIILGKLLFAPD 300
Qy 301 TPFTKRLMAQVNTFEELTLRLDRVREVMELGPRIFTFMNDSNVAMQLQWDEGRR 360
Db 301 TPFTKRLMAQVNTFEELTLRLDRVREVMELGPRIFTFMNDSNVAMQLQWDEGRR 360
Qy 361 QPRPGRDHMEALRSFLDPGSGYSQDAHADVGHVLTGRVTECLSDKLEAPSEAA 420
Db 361 QPRPGRDHMEALRSFLDPGSGYSQDAHADVGHVLTGRVTECLSDKLEAPSEAA 420
Qy 421 LVSRAQLLAEHRFWAGVFLGPDSDSDPTEHTPDLPDGPCHVRKIRMDIDVTRTNKIR 480
Db 421 LVSRAQLLAEHRFWAGVFLGPDSDSDPTEHTPDLPDGPCHVRKIRMDIDVTRTNKIR 480
Qy 481 DRFWDPGAADPLTLRYVMGFGFVYLQDLVERAAVRVLSGANPRAGLYLOQMPYCYVDD 540
Db 481 DRFWDPGAADPLTLRYVMGFGFVYLQDLVERAAVRVLSGANPRAGLYLOQMPYCYVDD 540
Qy 541 VFLRLVRSRLPLFLTLAWIYSVTLTKAVVREKTRLRDTRMRAGLSRAVLWGLWFLSCL 600
Db 541 VFLRLVRSRLPLFLTLAWIYSVTLTKAVVREKTRLRDTRMRAGLSRAVLWGLWFLSCL 600
Qy 601 GPFLLSAAALVVLKLGDIILPYSHPGVFLFAAFVATVTQSFLLSAFFSRANLAACG 660
Db 601 GPFLLSAAALVVLKLGDIILPYSHPGVFLFAAFVATVTQSFLLSAFFSRANLAACG 660
Qy 661 GLAYFSLYLPVYLCVAVWRDLRPAAGRVAAASLLSPVAFPGFCESIALLEQEGQAQWNV 720
Db 661 GLAYFSLYLPVYLCVAVWRDLRPAAGRVAAASLLSPVAFPGFCESIALLEQEGQAQWNV 720
Qy 721 TRPTADVFLSAQVSGLLLLDAALYGLATWYLVAVCPQYGIPEPWNPFRRSYWCGPRPP 780
Db 721 TRPTADVFLSAQVSGLLLLDAALYGLATWYLVAVCPQYGIPEPWNPFRRSYWCGPRPP 780
Qy 781 KSPAPCPTPLDPKVLVEAPPGSLSPGVSRSLKRPFGSPQALRGLSLDFYQGHITAF 840
Db 781 KSPAPCPTPLDPKVLVEAPPGSLSPGVSRSLKRPFGSPQALRGLSLDFYQGHITAF 840
Qy 841 GHNGAGKTTTSLISGLFPFGSGSAFTILGHDRSSMAAIPHLGVCQYQNVLPFMLTVDE 900
Db 841 GHNGAGKTTTSLISGLFPFGSGSAFTILGHDRSSMAAIPHLGVCQYQNVLPFMLTVDE 900

QY 901 HWFYGRKGLSAAVGPEDQLQDVGVLVSKQSQTQRLHSGMQKLSVAIAFVGGSQV 960
DB 901 HWFYGRKGLSAAVGPEDQLQDVGVLVSKQSQTQRLHSGMQKLSVAIAFVGGSQV 960
QY 961 VILDEPTAGVDPASRGITWELLKYREGTTLILSTHLLDEARLLGDRVAVVAGGLCCGC 1020
DB 961 VILDEPTAGVDPASRGITWELLKYREGTTLILSTHLLDEARLLGDRVAVVAGGLCCGC 1020
QY 1021 SPLFLRRHLGSGYILTLVKARLPLTNEKADTDMEGSDVTROEKNKGSGSRVGTPOLLA 1080
DB 1021 SPLFLRRHLGSGYILTLVKARLPLTNEKADTDMEGSDVTROEKNKGSGSRVGTPOLLA 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAELRLTGYISDTSLE 1140
DB 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAELRLTGYISDTSLE 1140
QY 1141 EIFLKVVECAADTDMEGSCGQHLCTGTAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
DB 1141 EIFLKVVECAADTDMEGSCGQHLCTGTAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRRRGLFAQIVLPALFVGLALVFSLIV 1260
DB 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRRRGLFAQIVLPALFVGLALVFSLIV 1260
QY 1261 PPGHYPALRLSPMYGAQVSFFSBDAPCDPGRARLLALQEALEPPVQVHSHRFS 1320
DB 1261 PPGHYPALRLSPMYGAQVSFFSBDAPCDPGRARLLALQEALEPPVQVHSHRFS 1320
QY 1321 PEVPAEAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
DB 1321 PEVPAEAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
QY 1381 NLTGRLNLSDFLVKTYPRLVROGLKTKKWNVEYGFSLGGRDPCLPQSGELGRSVEELW 1440
DB 1381 NLTGRLNLSDFLVKTYPRLVROGLKTKKWNVEYGFSLGGRDPCLPQSGELGRSVEELW 1440
QY 1441 ALLSPLPGCALDRVLKNTAWAHSUDAQDSLKIFNNKGWSMAFVNRASNAILRAHLP 1500
DB 1441 ALLSPLPGCALDRVLKNTAWAHSUDAQDSLKIFNNKGWSMTFVNRASNAILRAHLP 1500
QY 1501 PGPARHAISITLHNPLNITKEOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLEE 1560
DB 1501 PGPARHAISITLHNPLNITKEOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLEE 1560
QY 1561 RVTRAKHLQLMGGLSPITYLWLGNFVDMCNVLPACIVVLIFLAFQORAYVAPANLPALL 1620
DB 1561 RVTRAKHLQLMGGLSPITYLWLGNFVDMCNVLPACIVVLIFLAFQORAYVAPANLPALL 1620
QY 1621 LLLLLYGWSITPLMYPASFFSVPSYAYVVLTCINLFIGINGSMATFVLELFSQKQEV 1680
DB 1589 -----KLQEV 1593
QY 1681 SRILKQVFLIPHFCLGRGLDMVRNOAWADAFERLGRDQFQSPLRWEVGNKLLAMVIQ 1740
DB 1594 SRILKQVFLIPHFCLGRGLDMVRNOAWADAFERLGRDQFQSPLRWEVGNKLLAMVIQ 1653
QY 1741 GPLFLFTLLQHRSQLPQPRVRSPLPLGGEDEVDARERVRVQATQGDVLRNLTJK 1800
DB 1654 GPLFLFTLLQHRSQLPQPRVRSPLPLGGEDEVDARERVRVQATQGDVLRNLTJK 1713
QY 1801 VYRQMPAVDRCLCIGIPPGECFGLLVGNGAGKSTTFRMVGTDTLTSRGEAVLAGHSVAR 1860
DB 1714 VYRQMPAVDRCLCIGIPPGECFGLLVGNGAGKSTTFRMVGTDTLTSRGEAVLAGHSVAR 1773
QY 1861 EPSAAHLSMGVCPQSDAIFELLTGREHLELLARLKGVPPEAQVQATAGSGLARLGLSWAD 1920
DB 1774 EPSAAHLSMGVCPQSDAIFELLTGREHLELLARLKGVPPEAQVQATAGSGLARLGLSWAD 1833
QY 1921 RPAGTYSNGKRLATATALVGDPAVFLDPTTCMDPSARRFLWNLSLLAVVREGRSVNL 1980
DB 1834 RPAGTYSNGKRLATATALVGDPAVFLDPTTCMDPSARRFLWNLSLLAVVREGRSVNL 1893
QY 1981 TSHSMECEALCSRLAIAWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAFAFVAA 2040

DB 1894 TSHSMECEALCSRLAIAWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAFAFVAA 1953
QY 2041 BPPGSELREAHGGRRLRFQIPGGRCALARVFGELAVHGAHGVEDFSVSQTMLEEVFLYF 2100
DB 1954 BPPGSELREAHGGRRLRFQIPGGRCALARVFGELAVHGAHGVEDFSVSQTMLEEVFLYF 2013
QY 2101 SKDQKQKDEBTQEKAGVGVDPAPGLQHPKRVSQFLLDDPSPSTAETVL 2146
DB 2014 SKDQKQKDEBTQEKAGVGVDPAPGLQHPKRVSQFLLDDPSPSTAETVL 2059
RESULT 9
US-10-182-006-4
; Sequence 4, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: PCT/US01/02191
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1873
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-006-4
Query Match 84.9%; Score 9460.5; DB 15; Length 1873;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1835; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 MAFVTQMLLLKWNFMVRRRQPVQLLVELLWPLFFILVAVRHSHPLEHHECHFPNKP 60
DB 1 MAFVTQMLLLKWNFMVRRRQPVQLLVELLWPLFFILVAVRHSHPLEHHECHFPNKP 60
QY 61 LPSAGTVPWLQGLICNVNNTCFPQITPGEERGRISNFNDSLVSRLLADARTVLGASAH 120
DB 61 LPSAGTVPWLQGLICNVNNTCFPQITPGEERGRISNFNDSLVSRLLADARTVLGASAH 120
QY 121 TLAGLGKLIATLRAARSTAOPTKQSPLEBPPMLDVAELLTSLLRTEISGLALGOQBPL 180
DB 121 TLAGLGKLIATLRAARSTAOPTKQSPLEBPPMLDVAELLTSLLRTEISGLALGOQBPL 180
QY 181 HSLEAAEDLAQELALRLSLVELRALLQRPRTSGPLELLSEALCSVRGSPSTVGPSLW 240
DB 181 HSLEAAEDLAQELALRLSLVELRALLQRPRTSGPLELLSEALCSVRGSPSTVGPSLW 240
QY 241 YEASDLMELVQBPESALPDSSLSPPACSELICGALDHPHLSRLLRLLKPLILGLLPAD 300
DB 241 YEASDLMELVQBPESALPDSSLSPPACSELICGALDHPHLSRLLRLLKPLILGLLPAD 300
QY 301 TPFTKRLMAQVNRITFEELTLRLDVRVWEMLGPRIFTEPMNDSNVNMLQRLLOMDEGR 360
DB 301 TPFTKRLMAQVNRITFEELTLRLDVRVWEMLGPRIFTEPMNDSNVNMLQRLLOMDEGR 360
QY 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
DB 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
QY 421 LVSRALQALLAHRFWAGVFLGPDSSDPTCHPTDLPFGPHVRKIRMDIDVVTNRKIR 480
DB 421 LVSRALQALLAHRFWAGVFLGPDSSDPTCHPTDLPFGPHVRKIRMDIDVVTNRKIR 480

```
QY 481 DRFDWPGPADPLTDLYRVWGGFYVLQDLVERAAVRVLSGANPRAGLYLOOMPYPCTYDD 540
Db 481 DRFDWPGPADPLTDLYRVWGGFYVLQDLVERAAVRVLSGANPRAGLYLOOMPYPCTYDD 540
QY 541 VFLRVLSRSLPLFTLTLAWISVTLTKAVVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
Db 541 VFLRVLSRSLPLFTLTLAWISVTLTKAVVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
QY 601 GPFLLSAAALVVLKLGDIILPYSHPGVVFLLAFAVAVTQSFLLSAFFSRANLAAACG 660
Db 601 GPFLLSAAALVVLKLGDIILPYSHPGVVFLLAFAVAVTQSFLLSAFFSRANLAAACG 660
QY 661 GLAYFSIYLPIVLCVANDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGQAHNVG 720
Db 661 GLAYFSIYLPIVLCVANDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGQAHNVG 720
QY 721 TRPTADVFLSAQVSGILLDDAALYGLATWLEAVCPQOYGIPEPWNPFRRSVCWGRPPP 780
Db 721 TRPTADVFLSAQVSGILLDDAALYGLATWLEAVCPQOYGIPEPWNPFRRSVCWGRPPP 780
QY 781 KSPAPCPTPLDPKVLVEEAPPGVSPGVSVRSLKRPFGSPQALRGLSLDFYQGHITAFI 840
Db 781 KSPAPCPTPLDPKVLVEEAPPGVSPGVSVRSLKRPFGSPQALRGLSLDFYQGHITAFI 840
QY 841 GHNGAGTKTTLISLSGLFPSPGGSAPFILDHVRSSMAAIRPHLGVCPOYNNVLPDMLTVDE 900
Db 841 GHNGAGTKTTLISLSGLFPSPGGSAPFILDHVRSSMAAIRPHLGVCPOYNNVLPDMLTVDE 900
QY 901 HVWFYGRKGLSAAVGPEDRLLDQVGLVSKQSVQTRHLSSGMQRKLSVAIAFVGGSQV 960
Db 901 HVWFYGRKGLSAAVGPEDRLLDQVGLVSKQSVQTRHLSSGMQRKLSVAIAFVGGSQV 960
QY 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLIISTHLDDEALIGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLIISTHLDDEALIGDRVAVVAGRLCCCG 1020
QY 1021 SPLFLRHLSGGYVLTIVKARLPLTTNEKADTMEGSDVDTROEKKNGSQSRVCTPOLL 1080
Db 1021 SPLFLRHLSGGYVLTIVKARLPLTTNEKADTMEGSDVDTROEKKNGSQSRVCTPOLL 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGANDGSPATLFRELDTRLAELRLTGYISDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLPYTGANDGSPATLFRELDTRLAELRLTGYISDTSLE 1140
QY 1141 EIFLKVVVECAADTDMEDGCGQLCTGIAGLDVTLRLKMPQPTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVVECAADTDMEDGCGQLCTGIAGLDVTLRLKMPQPTALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARESRGLFAQIVLPALFVGLALVFSLIY 1260
Db 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARESRGLFAQIVLPALFVGLALVFSLIY 1260
QY 1261 PPFQHYPALRLSPMYGAQVFFSDEDPGDPGRARLLEALLQEAAGLEPPVQHSSHRFSA 1320
Db 1261 PPFQHYPALRLSPMYGAQVFFSDEDPGDPGRARLLEALLQEAAGLEPPVQHSSHRFSA 1320
QY 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARELLPDCPAAAGPPPPQAVTSGSRVQ 1380
Db 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARELLPDCPAAAGPPPPQAVTSGSRVQ 1380
QY 1381 NLTCRNLSDFLVKTYPRIVROGLTKKWVNEVRVGGFSLGGRDGLPSGOELGRSVEELW 1440
Db 1381 NLTCRNLSDFLVKTYPRIVROGLTKKWVNEVRVGGFSLGGRDGLPSGOELGRSVEELW 1440
QY 1441 ALLSPLPGALDRVLKNTLAWHSLDAQDSLSKIWFNNKGWHSMAFVFNRAASNATILRAHLP 1500
Db 1441 ALLSPLPGALDRVLKNTLAWHSLDAQDSLSKIWFNNKGWHSMAFVFNRAASNATILRAHLP 1500
QY 1501 PGPARHAHSITTLNHPNLNLTKEQISEAALMASSVDVLVSIQVFPAMSFVPASFTLVLEE 1560
Db 1501 PGPARHAHSITTLNHPNLNLTKEQISEAALMASSVDVLVSIQVFPAMSFVPASFTLVLEE 1560
QY 1561 RVTRAKHLQMLGSLPTLYWLGFLWDMCNVLPACIVVLIFLAQORAYVAPANLPALL 1620
```

```
Db 1561 RVTRAKHLQMLGSLPTLYWLGFLWDMCNVLPACIVVLIFLAQORAYVAPANLPALL 1620
QY 1621 LLLLLXGWSITPLMWYPASFFSVSPSTAYVYVLTICINLFIGINGSMATFVLBSFSDQKLOEV 1680
Db 1621 LLLLLXGWSITPLMWYPASFFSVSPSTAYVYVLTICINLFIGINGSMATFVLBSFSDQKLOEV 1680
QY 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVGKLLAMVIO 1740
Db 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVGKLLAMVIO 1740
QY 1741 GPLFLFTLLQHRSQLLPQVRSLPLLEEEDVVARERERVVQATQGDVLRNLTK 1800
Db 1741 GPLFLFTLLQHRSQLLPQVRSLPLLEEEDVVARERERVVQATQGDVLRNLTK 1800
QY 1801 VYRGORMPAVDRICLGIIPGCEGCLLGVNGAGTKTSTFRMVTGTLASRGAVALAGHSVAR 1860
Db 1801 VYRGORMPAVDRICLGIIPGCEGCLLGVNGAGTKTSTFRMVTGTLASRGAVALAGHSVAR 1860
QY 1861 EPSAAHLSMG-YCQSDAIFELL 1882
Db 1855 -----HLLLTGYCM---PIFVLL 1869

RESULT 10
US-09-995-542-2
; Sequence 2, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Uliar, Jaarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2167
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-995-542-2

Query Match 77.3%; Score 8613.5; DB 9; Length 2167;
Best Local Similarity 76.7%; Pred. No. 0;
Matches 1678; Conservative 163; Mismatches 283; Indels 65; Gaps 11;

QY 1 MAFWTQMLLLWKNFMVRRRQPVQLLVELLWPLFLPFILVAVRHSHPLEHHECHFPNKP 60
Db 1 MAFCTQLMLLLWKNYTYRRRQPIQLLVELLWPLFLPFILVAVRHSHPLEHHECHFPNKP 60
QY 61 LPSAGTVPMVLOGLICNVNNTCFQLTGPEEPGRLSNFNDLSILVSRLLADARTVLGSAHR 120
Db 61 LPSAGTVPMVLOGLICNVNNTCFQLTGPEEPGRLSNFNDLSILVSRLLADARTVLGSAHR 120
QY 121 TLAGLGLKLIATLRAARSTAQP---QPTKQSPLEPPMLDVAELTSLRRTESLGLAQ 176
Db 121 MLDALGKLIPLVAVGGGARQESDQPTSQ-----SVTKLEKILQRASLDPVLGQA 173
QY 177 QEPHLSLLEAAEDLAQELLALRSILVELRALLQRPRTSGPGLLELLSEALCSVRGSSVGP 236
Db 174 QDSMRKFSDAIRDLAQELLTLPSLMELRALLRRPRGSAGSLELVSEALCSVTGKSPSGGL 233
QY 237 SLNWYEASDIWELVQGPESALPDSSUSPACSELIGALDSHPILSRLLRRLKPLILGKLL 296
Db 234 SLNWYEANQNEFMGPEVAPALPDNSLSPACSEFVGLTDDHPVSRLLRRLKPLILGKIL 293
QY 297 FAPDTPTRKILMAQVNRFTFEELTLDRDVRVWEMGLPRIFTFMNDSSNVAMQLLOQMD 356
Db 294 FAPDTPTRKILMAQVNRFTFEELTLDRDVRVWEMGLPRIFTFMNDSSNVAMQLLOQMD 353
```


US-09-995-542-3

Query Match 75.3%; Score 8389.5; DB 9; Length 2121;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;

Qy 47 PPLEHHECHFPNKPLPSAGTVPWLOGLICNVNNTCFPOLTPGEPGBLSFNDSLSVRL 106
Db 1 PPLEHHECHFPNKPLPSAGTVPWLOGLICNVNNTCFPOLTPGEPGBLSFNDSLSVRL 60

Qy 107 ADARTVLGGASARTLGLKLIATRAARSTAP---OPTKQSPLEPPMLDVAEELTS 162
Db 61 ADRTVILGHSIQMLDALGKLIIVLRAVGGARPQESDQTSQ-----SVTKLEK 113

Qy 163 LLRTESLGLAQOAEPLHSLEAEEDLAQELLALRLSLVELRALLQRPRTSGPLELSE 222
Db 114 ILQASLDPVLQAOQDMRKFSDAIRDAQELLTLPSLMELRALLRRPRGSAGSLELVSE 173

Qy 223 ALCSVRGSSSTVGPSLWNYEASDLMELVQEPESALPDSSLSPACSELIGALDSHPLSRL 282
Db 174 ALCSVTGPPSGGLSLWNYEANOEFMGPEVAPALPDNSLPACSEFVGLTDDHPVSRL 233

Qy 283 LWRRLKPLILGKLFAPDTPTRKMAQVNTPEELTLRDVREVWMLGPRIPTFNWDS 342
Db 234 LWRRLKPLILGKLIIFAPDTPTRKMAQVNTPEELTLRDVREVWMLGPRIPTFNWDS 293

Qy 343 SNVAMLQRLQOMDEGRQPRPGGRDHMEALRSFLDPGSGYSQDADAHVGLVGTGR 402
Db 294 TNVAMLQRLQOMDEGRQPRPGGRDHMEALRSFLDPGSGYSQDADAHVGLVGTGR 353

Qy 403 VTECLSDKLEAESEALVSRLQALLAEHRFWAGVVPFLGPESSDTEHPTDLPGRHV 462
Db 354 NMCEVSLDKLEAESEALVSRLQALLAEHRFWAGVVPFLGPESSDTEHPTDLPGRHV 413

Qy 463 RIKRMDIDVTRNKIRDFWDPGAADPTDLRYVWGGFVYLQDLVERAAVRLSGAN 522
Db 414 RFIKRMIDVTRNKIRDFWDPGAADPTDLRYVWGGFVYLQDLVERAAVRLSGAN 473

Qy 523 PRAGLYIQMPYPCYVDDVFLRVLRSPLPLTLIAWYSVTLTKAVVREKETRLRDMR 582
Db 474 SRTGLYIQMPYPCYVDDVFLRVLRSPLPLTLIAWYSVTLTKAVVREKETRLRDMR 533

Qy 583 AMGLSRVILWGLFSLCLGPELSSAALLVILVGLDILPYSHPGVFLFLAAFAVATVQ 642
Db 534 AMGLSRVILWGLFSLCLGPELSSAALLVILVGLDILPYSHPGVFLFLAAFAVATVQ 593

Qy 643 SFLLSAFFSRANLAAACGLAYFSLPYVLCVAVWRDLRPAAGRVAAASLLSPVAFGCE 702
Db 594 SFLLSAFFSRANLAAACGLAYFSLPYVLCVAVWRDLRPAAGRVAAASLLSPVAFGCE 653

Qy 703 SLALLEEQGGAQWHNVGTRPTADVFLAQVSGLLLLDAAALYGLATWYEAVCPCQYQIP 762
Db 654 SLALLEEQGGAQWHNVGTRPTADVFLAQVSGLLLLDAAALYGLATWYEAVCPCQYQIP 713

Qy 763 EPWNFPRRSYWCGRPKSPAPCTPLDPKVLVEEAPPGLSPGVSRLEKRPSPGP 822
Db 714 EPWNFPRRSYWCGRPKSPAPCTPLDPKVLVEEAPPGLSPGVSRLEKRPSPGP 773

Qy 823 ALRGLSLDFYOGHTITAFILHNGAGKTTLSILSLGFLPPSGGSASFILGHVDRSSMAAIRPH 882
Db 774 ALQGLNLDYEGHITAFILHNGAGKTTLSILSLGFLPPSGGSASFILGHVDRSSMAAIRPH 833

Qy 883 LGVCPQNVLPDMULTVDEHVMFYGRLLKGLSAAVVGPEQDRLLQDVGLVSKQSVTRHLSG 942
Db 834 LGICPQNVLPDMULTVDEHVMFYGRLLKGLSAAVVGPEQDRLLQDVGLVSKQSVTRHLSG 893

Qy 943 GMQRKLSVAIAFVCGSQWVILDEPTAGVDPASRREGIWEILLKYREGTLLISLTHLDEAE 1002
Db 894 GMQRKLSVAIAFVCGSQWVILDEPTAGVDPASRREGIWEILLKYREGTLLISLTHLDEAE 953

Qy 1003 LLGDRVAVAGRLCCCGSPFLRRHLGSGYLLTVKARPLPTTNE-KADTDMEGSVDTTR 1061
Db 954 LLGDRVAVAGRLCCCGSPFLRRHLGSGYLLTVKSSQSLVTHDAKGDS-----DPR 1008

Qy 1062 QEKNGSQGS-----RVGTPOLLALVQHWTPGARLVEELP 1096
Db 1009 REKSDNGRSTDAFTRGTSKSNQAPAGVPIPTSTARIILELVQHQHVPGAQVDELPL 1068

Qy 1097 HELVLVLPYTGADGSPATLFRLELDRLAELRTLGTGISTDLEEFILKVVBECAADTDM 1156
Db 1069 HELLLVLPYAGADGSPAMVQFQELDQOELLGLTGTGISTDLEEFILKVVBECAADTDM 1126

Qy 1157 EDGSCGHLCTGAGLDVTLRLKMPQOETALENGEPAGSA--PETDQSGSPDVGVRVQGG 1214
Db 1127 GDSRPQLHRT-----CTPQPTGPEASVLENGELAKLVLDPAQOGLAPNA-AQVOGG 1179

Qy 1215 ALTRQLOQALLKRLARRSRRLFAQIVLPALFVGLALVSLIIVPPFGHYPALRLSPT 1274
Db 1180 TLTCQQLRALLHRLARRSRRLFAQIVLPALFVGLALVSLIIVPPFGHYPALRLSPT 1239

Qy 1275 MYGAQVSFFSEADAGDPGRARLLEALLQEAAGLEPPVQHSR-----FSAPEV 1323
Db 1240 MYGQVSFFSEADAGDPGRARLLEALLQEAAGLEPPVQHSR-----FSAPEV 1299

Qy 1324 PAEVAKYLASGNWTPSPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGSEVWQNL 1383
Db 1300 PPDVASILASGNWTPSPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGSEVWQNL 1359

Qy 1384 GRNLSDFLVKTYPRLVROGLKTKWNEVRYGFSLGGRDPGLPSQELGRSEELWALL 1443
Db 1360 GRNVSDFLVKTYPSLVRRGLKTKWNEVRYGFSLGGRDPGLPSQELGRSEELWALL 1419

Qy 1444 SPLPGGALDRVLKNTLTAWAHSLDAQSLKTFWKNKWHSMVAFVNRASNAILRAHLPPGP 1503
Db 1420 SPOGNALDRILNLTQWALGLDARNSLKTFWKNKWHSMVAFVNRANGLLHALLPSGP 1479

Qy 1504 ARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDLVSI CVVFAMS FVPASFTLVLIBERT 1563
Db 1480 VRHAHSITTLNHPNLNLTKEQLSEATLIASSVDLVSI CVVFAMS FVPASFTLVLIBERT 1539

Qy 1564 RAKHLQMLGSLSTLWLNFLMDMNYLPACIVLILFLAFOORAYVAPANLPAILLIL 1623
Db 1540 RAKHLQMLGSLSTLWLNFLMDMNYLPACIVLILFLAFOORAYVAPANLPAILLIL 1599

Qy 1624 LLYGWSITPLMYSPASFFSVPTAYVVLTCINLFIGINGSMATFVLELSDQKLOEVSRI 1683
Db 1600 LLYGWSITPLMYSPASFFSVPTAYVVLTCINLFIGINGSMATFVLELSDQKLOEVSRI 1659

Qy 1684 LKQVFLIPPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWEVVKNNLLAMVIOGPL 1743
Db 1660 LKQVFLIPPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWDIIGKNLLAMMAQGPL 1719

Qy 1744 FLIPLTLLQHRSQLPQPRVRSIPLIGEEDVARERERVQCATQGDVLVRLNLTKVYR 1803
Db 1720 FLIPLTLLQHRSQLPQPRVRSIPLIGEEDVARERERVQCATQGDVLVRLNLTKVYR 1779

Qy 1804 GQMPAVDRILCIGIPGCECFGLLVNGAGKTSFRMTVGTDLASRGAVALAGHSVAREPS 1863
Db 1780 GQMPAVDRILCIGIPGCECFGLLVNGAGKTSFRMTVGTDLASRGAVALAGHSVAREPS 1839

Qy 1864 AAHLSMGYCQSDAIPELLTGREHLELALRLGRVPEAQVAQTAGSGLARLGLSWYADPA 1923
Db 1840 AAHLSMGYCQSDAIPDLTGREHLELALRLGRVPEAQVAQTAGSGLARLGLSWYADPA 1899

Qy 1924 GTYSGGNKRLATALLVGDPAVVFDEPTTGMDDPSARRFLMNSLLSVVREGSVMLTSH 1983
Db 1900 GTYSGGNKRLATALLVGDPAVVFDEPTTGMDDPSARRFLMNSLLSVVREGSVMLTSH 1959

Qy 1984 SMECEALCSRLAIMVNGRPRCLGSPHOLKGRFAAGHTLTLRVPAAASQPAAPAAFEFF 2043
Db 1960 SMECEALCSRLAIMVNGRPRCLGSPHOLKGRFAAGHTLTLRVPAAASQPAAPAAFEFF 2019

Qy 2044 GSLELREAHGRLRFPQPPGRCALARVFGELVHGAHGVEDFSVSTQMLLEEVFLYFSKD 2103
Db 2020 GAELREHVGSRILRFPQPPGRCALARVFGELVHGAHGVEDFSVSTQMLLEEVFLYFSKD 2079

Db 1907 QONDIIEIKELTKIYRRKRKPAVDRI CVGIPFCECFGLLVGNAGKSGSTFRKMLTGDITVT 1966
QY 1848 RGEAVLAGHSVAREPSAHLSCGYCQSDAIPELLTGRHLELLAARLCRVPEAQATAG 1907
Db 1967 RGAFLNKSILSNHEVQNGYCDQFOAI TELTGRHVEFFALLRGVPEKVGKE 2026
QY 1908 SGLARLGWSYADRPA GTYSGGNKRKLATALAVGDPAPVVFDEPTTGMDSARRPLMNS 1967
Db 2027 MAIRKGLGVYKEVAGNYSNGNKRKLSTAMALIGGPPVVVFDEPTTGMDSARRPLMNC 2086
QY 1968 LLAVYREGSVMLTSHSMECEALCSRLAIMNGRFRCLGSOHLKGRPAAGHTITLRVP 2027
Db 2087 ALSVKEGSGSVLTSHSMECEALCTRMALMNGRFRCLGSOHLKGRFGDGTIVVRIA 2146
QY 2028 AARS--QAAAAFAAEFFPGSELREAHGGRRLRFQLPFGRCALARVFGELAVHGAHVVED 2085
Db 2147 GSNPDLKPQDFGLAFPGSVLKEHRNMLQVL--PSSLSLSIARIFSILSOSKKRLHIED 2205
QY 2086 FVSQTMLEEVFLYSKQKQDB---DTBEOKEAGVGVDPAFGLQHPKRVSOFLDDPSTA 2142
Db 2206 YSVSQTLDDQVFNFAKQSDDDHLKDLSLHKNQTV--VDVAV-----LTSFLQDEKVK 2257
QY 2143 ET 2144
Db 2258 ES 2259
RESULT 14
US-10-313-641-10
; Sequence 10, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P023151052
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-313-641-10
Query Match 51.8%; Score 5768.5; DB 14; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;
QY 1 MAFWTQMLLWKNWYRRQPVOLLVLLWPLFLFFILVAVRSHSPLEHHECHFPNKP 60
Db 1 MACWPQLRLWKNLTFRRRTQCLLLEWAMPFLFILISVRUSYPYPYQHECHFPNKA 60
QY 61 LPSAGTVPWLOGLICNVNNTCFPQLTGPEEGRISNFNDLSVSRLLADARTVLGSAHR 120
Db 61 MPSAGTLFWQGIICNANNPCRYPTGAEAFGVGNFNKSVIARLFSADARLLLYSQDXT 120
QY 121 TLAGUGKIATUARAARSTAQ-----POPTKQSPLEPML-- 154
Db 121 SMKMRKVLRTLQIICKSSNLKLDVLDNFTSGFLYHNLSPKSTVDKMLRADVLH 180
QY 155 -----DVNEL-----LTSLL 164
Db 181 KVFLQGYQLHTSLCNGSKSEMIQLGQGVSELGCLPREKLAARVLRNSMDILKPI 240

QY 165 RT--BSLGLALQAOEPLHLSLEAAEDLAQELLARLSLVELR---ALLQRPRTSGPLEL 219
Db 241 RFLNSTSPPSKELAEATKLLHSLGTLAQELFSRMSDMRQEVMLFNWSSSSTQI 300
QY 220 ---LSEALCSVRGPGSTVGPSLNWYEASDLMBELVG---QBPESALPDSLSGPACSELIG 272
Db 301 YQAVSRI VCGHPGEGGLKIKSLNWEYEDNNYKALFGNGTEEDAETFDNSTTTPYCNCLMK 360
QY 273 ALDSPLSRLRLRRLKLLGLKLLPAPDTPFKLMAQVNRTEELTLARDVREVWMLG 332
Db 361 NLESPLSRIIWALKPLLVGLKILYTPDTPATQVMAEVNKTQELAVFHDLEGWMEELS 420
QY 333 PRIFTFMDSSNVAMLQRLQWQDEGR--RQRPGGRD-----HMEALRSFLDP 379
Db 421 PKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTQADIVAFLAGHPEDVQS--- 476
QY 380 SGGG--YSWQDAHADVGHVGLTRVTECLSLDKLEAAPSEAAALVSRLAQLAEHRFWAG 437
Db 477 -SNGSVYTRAEAFNETNOAIRTISRMECVNLKLEPIATEVWLINKSMELDERKFWAG 535
QY 438 VVFLGPESSDPTTEHTPDGLGPHVRIKIRMDIDVVTNRNKIRDFWDPGPAADPLTDLR 497
Db 536 IVFTGITPGSIELPH-----HVYKIRMDIDNVERTNKIKGVTWDPGPAADPEDNR 587
QY 498 YWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPICYVDVDFLRLVLSRLPLFLTLA 557
Db 588 YWGGFAYLQDVVEQAIIRVLGTGTEKTVGYMQMPYCYVDDIIFLRVMSRSMPLFMTILA 647
QY 558 WIYSVTLTVKAVVREKETRLRDTWRAMGLSRAVLWGLWFLSCGLGPFLLSAALLVLVLKLG 617
Db 648 WIYSVAVTIKGIYVEKEARLKETWRIMGLDLSILWFSWFISSILPLLVSAGLLVLVLKLG 707
QY 618 DILYSHGPGVVFELAAEAVATVTSFILSAFFSRLANAAACGLAYSLYLYPYVLCVAV 677
Db 708 NLLPYSDPSVVFVLSVFAVVTILQCLISTLFSANLAAACGGIYFTLYPYVLCVAV 767
QY 678 RDRLPAGGRVAAASLLSPVAFGCGESLALLBEEQGEAGVHNVGTRPT--ADVFLAQVSGL 736
Db 768 QDYVGTFLKI PASLLSPVAFGCGEYFALFEEQIGVQWQDNLFESFVEEDGNLTSVSM 827
QY 737 LLLDAAALYGLATWYLEAVCPQGYGIPPEWNPFFRRSYWCGRRPPKSPACPTPLD--PKVL 795
Db 828 MLFDTFLYGVMTWYIEAVFPQGYGIPRPWYFPCTKSYWFGESDESKHSPGSKRISBIC 887
QY 796 VEEAPPGLSPGVSVSRLEKRPGPSQPALRGLDFOGHITAFILGHNGAGKTTTILSILS 855
Db 888 MEEEPHTKLKVSQNLVKYRDGMKAVDGLALNFYEGQITSFILGHNGAGKTTTWSILT 947
QY 856 GLFPFPGGSAPILGHDVRSSMAAIRPHLGVCPCQYNVLFDMLTVDHVMFYGRKLGLSAAV 915
Db 948 GLFPPTSGTAYILGKD IRSEMTIRQNLGVCPQHNVLFDMLTVEEHIWIFYARLKGLSEKH 1007
QY 916 VGPQDRLLQDVGL--VSKQSVQTRHLSGGMQRKLSVAIFVGGSOVVIDEPTAGVDPAS 974
Db 1008 VKAEMEQNALDVLGPPSSKLSKTSQLSGGMQRKLSVALAFVGGSKVVIDEPTAGVDPS 1067
QY 975 RRGWELLKYREGRTTILSTHLLDEALLGDRVAVAGGLCCCGSPILFRHRLGSGVY 1034
Db 1068 RRGWELLKYRQRTIILSTHMDDEADVLDGRITAIISHGKLCCCGSSFLKNQKGTGY 1127
QY 1035 LTLVKARLPLTN-----EKADTMEGSDVTRQEKNGSQSGSVGTPTQALLAV 1082
Db 1128 LTLVKQDVESLSSCRNSSSTSVYLLKEDSVSQSSSDAGLGDHSDTLTIDVSAISNLI 1187
QY 1083 QHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRLEDTRIAELURLYGYGISDTSLEBI 1142
Db 1188 RKVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRLSLDLSIGSYSETTLEBI 1247
QY 1143 FLKVVEECAADTMDGSCGOHLCTGIAGLDVTLRLKMPPOETALE--NGEPAGSAPTD 1200
Db 1248 FLKVAESGVDAETSDGTLPARNRRAFG--DKQSCLRPFTEBDDAADPNDSDIDPSRETD 1306
QY 1201 QGSGPDVAG--RVQGWALTQQLALLKRLFLARRRRGLFAQIVLPALFVGLALVFSL 1258

```
Db 1307 LLSGMDGKGYQVKGWKLTKQQFVALLWKKLLIARRRKGFFAQIVLPAVFCIALVFSL 1366
Qy 1259 IVPFPGHYPALRLSPMYGAQVSFSDAPCDPGRRARLLALQEAQ-----LEE 1308
Db 1367 IVPFPGKYPSELPQWYNEQYTFVSNDAPEDTCTLELLNALTKDPCGFRCMEGNFI 1426
Qy 1309 PPVQSHSRFSAPBPAEVAKVLAGNWTPESSPACQSQPGARRLLPCDPAAGGPPP 1368
Db 1427 TPCQAGEEWTAPVQTIIMDLFQNGNWTQNPSPACQSSDKIKMLPVCPPGAGLPP 1486
Qy 1369 PAVTGSSEVQNLTKGNLSDFLVKTPRLVROGLTKKWNVEVYGGFSLG-GRDGLP 1427
Db 1487 PQRKQNTADILQDITGNISDYLVTQVLIQIATSKKNKIWNNEFRYGGFSLGVSNTQALP 1546
Qy 1428 SGQELGRSVBELWALLSPFGGALDRVLKMLTAHSLDAQDSLKIFWNNKGWHSMAVF 1487
Db 1547 PSQEVNDIAIKOMKHLKLAKDSSADREFLSNGREMTGLDTKNVVKWVFNKGWHAISFL 1606
Qy 1488 NRASNAILRALHPLPGPARHAHSITLNLHPLNLTKQELSEALMASSVDVLVSVICVPAMS 1547
Db 1607 NVINNAILRANLQGENSPHYGITAFNHPNLTKQQLSEVALMTTSVDVLVSVICVPAMS 1666
Qy 1548 FVPASFTLVLEERTVRAKHLQMGGLSPTYLWLNGLNFWDMCNVLPACIVVLIFLAFQ 1607
Db 1667 FVPASFVFLIOERVSXAKHLQFTSGVKPVLYLSNFWDMCNVLPATVILIFICFQQ 1726
Qy 1608 RAYVAPANLPALELLLYLWGSITPLMPASFFSVSPSTAYVVLTCINLFIGINGSMATF 1667
Db 1727 KSYVSTNLPLVALLELLLYLWGSITPLMPASFVKIPSTAYVVLTSVNLFIGINGSVATF 1786
Qy 1668 VLELFSOKIQEVSRIILKQVFLIPPHFCGLRGLIDMVYNOAMADAFERLGRDQFQSPLRW 1727
Db 1787 VLELFTDKLNNINDILKSFLVLPHPFCGLRGLIDMVYNOAMADALERFENRFPVPSLW 1846
Qy 1728 EVGKNNLAWIQGLFLLFTLLQHSQLLPOPRVRSPLGLGEDEEDVARERERVQGA 1787
Db 1847 DLVRNLFMAVEGVVFLITVLIOYRFFRPRVNAKLSPLNDEEDVRRERQRILGG 1906
Qy 1788 TQGDVLVRLNLTKYVQRMPADVRLCLGIPPGECFGLLVNGAGKSTTFRMVGTDTIAS 1847
Db 1907 QONDILSEIKELTKYRKRKPAVDRI CVGIPPGECFGLLVNGAGKSTFKMLTGDITVT 1966
Qy 1848 RGEAVLAGHSVAREPSAAHLSMGYCPSQDAIFELLTGREHLELLARLGRVPEAQVATAG 1907
Db 1967 RGDAPLKNLSILNIHVHQNMGYCPQFDAITELLTGREHVEFPALLRGVPEKEVGKVE 2026
Qy 1908 SGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVVPFLDBPTTGMDPSARRFLWNS 1967
Db 2027 WAIKJGLVYKEGYAGNYSGNKRKLSTWALIGGPVFLDBPTTGMDPKARFLWNC 2086
Qy 1968 LLAVVREGRSYMLTSHSMECEALCSRLAIVNGRFRCLGSPQLHKGKFAAGHTLTLRVP 2027
Db 2087 ALSVVKEGRSVLTSMSMECEALCTRMAIVNGRFRCLGSPQLHKNRFGDGYTIVVRIA 2146
Qy 2028 AARS--OPAAFAVAEPFGSELREAGGRRLRFQLPPGRCALARVFGELAVHGAHEVED 2085
Db 2147 GSNPDLPVQDFFGLAFPGSLVKEKHNMLQYQL-PSLSLSLARIFSILSQSKRLHIED 2205
Qy 2086 FSVSQTMLEEVLYFSKQKQDE--DTEQKEAGVGVDPAPGLQHPKRVSOFLDDPSTA 2142
Db 2206 YSVSQTTLLDQVFNFAKQSDDDHLKOLSLHKNQTV-VDAV-----LTSFLQDEKVK 2257
Qy 2143 ET 2144
Db 2258 ES 2259
```

RESULT 15

US-10-428-551-9

; Sequence 9, Application US/10428551

; Publication No. US20030229062A1

; GENERAL INFORMATION.

```
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P02351US3
; CURRENT APPLICATION NUMBER: US/10/428,551
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
; US-10-428-551-9
```

```
Query Match 51.8%; Score 5768.5; DB 14; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;
```

```
Qy 1 MATWTQMLLLWKNFVRRRQPVQLLVELLWPLFLFVLVAVRHSHPHHECHFPNKP 60
Db 1 MACPQRLLLWKNLFRRRQTCQLLLEVAVPLFIILISVRLSVPPYEQHECHFPNKA 60
Qy 61 LPSAGTVPMLOGLICNVNNTCPQLTPGEBPGRLSNFNDLSVRLKLADARTVLGGASAH 120
Db 61 MPSAGTLPWVGIIICNANNPCFRYPPTGEPAGVGVGNFNKSIVARLFSADARLLYSQKT 120
Qy 121 TLAGLGLKLTATRAASTAQ-----POPTKOSPLEPPML-- 154
Db 121 SMKDMRKVLTLQAIKKSSSNLQLQFLVDNETFSGFLYHNLSLPKSTVDKMLRADVLH 180
Qy 155 -----DVAEL-----LTSLL 164
Db 181 KVFLOGYQLHLTSLCNGSKSEMIQIQDQEVSLCGLPREKLAARVLRNSMDILKPI 240
Qy 165 RT--ESLGLAQAQAPLHSLLEAAEDLAQELLALRSLVELR---ALLQPRGTSGPLE 219
Db 241 RTLNSISPPSKELAEATKLLHSLGTLAGELFSMRSDMRQEVMPFLTVNSSSSSTQI 300
Qy 220 ---LSBALCSVRGSPSTVSGSLNWNVYASDLMELVG---QEPESALPDSSLSPACSELIG 272
Db 301 YQAVSRIVCGHPGEGGLIKIKLNWYEDNNYKALFGNGTEDEAETFYDNTSTTTCNDLMK 360
Qy 273 ALDSHPLRLLWRRLKPLILGLKLLPAPDTPFTRKLAQVNRNTEELTLDRDREVWELG 332
Db 361 NLESSPLSRIIWKALPPLLVGKILYTPDTPATQVMAEVNKTFOELAVFHDLEGWBEELS 420
Qy 333 PRIFTFMDSSNVAMQLQRLQLOMDEGR-ROPRPGGRD-----HMEALRSFLDP 379
Db 421 PKIWTFWNSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTADQDIVAFLAKHPEDVQS--- 476
Qy 380 GSGG--YSWODAHADVGHVLTIGRVVTECLSLDKLEAAPSEAAALVSALQALLAHRFWAG 437
Db 477 -SNGSVYTWREAFTNETQAIRTIISRFMECVNLKLEPIATEVWLNKSMELLDERKEFWAG 535
Qy 438 VVFLGEDSSDPTHEHFPDGLPGHVRKIRMDIDVYTRTNKIRDRFWDGCPAADPLDLR 497
Db 536 IVFTGITPGSIBLPH-----HVYKIKRMDIDNVERTNKIKDGYWDGPGPRADPFEDMR 587
Qy 498 YWNGGFVYLQDLVERAAVRVLSGANPRAGLYLOQMPYPCYVDDVFLRVLSRSLPLFTLA 557
Db 588 YWNGFAYLQDVVEQAIIRVLVTGTEKTYVMQMPYPCYVDDIFLRVMSRSMPLFWTLA 647
Qy 558 WIYSVTLTVKAVVREKTRURDTRWAMGLRAVLWLGWFTSLCGLPFLLSAALLVLVLKLG 617
Db 648 WIYSVAVIIGKIVVEKEARKETWIRIMGLDNLISLWFSWFISSLTPLIVSAGLLVILKLG 707
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 180.939 Seconds
(without alignments)
4254.650 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMYRR.....QHPKRVQFLDDPSTAETVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11143	100.0	2146	6	Abg72696 Human ATP
2	11130	99.9	2146	4	Aau04483 Human PD-
3	11130	99.9	2180	5	Aao14210 Human tra
4	11128	99.9	2146	5	Abp52096 Homo sapi
5	11128	99.9	2146	6	Abu08466 Human ABC
6	11107	99.7	2144	5	Aau09174 Human tra
7	11107	99.7	2144	7	Add37429 Human tra
8	10639.5	95.5	2059	6	Abu54629 Human NOV
9	10264	92.1	2008	6	Abu08464 Amino aci
10	10171.5	91.3	1993	6	Abu08465 Amino aci
11	9460.5	84.9	1873	4	Aau04484 Human PD-
12	8613.5	77.3	2167	6	Abg72695 Mouse ATP
13	7596	68.2	1550	6	Abg72697 Human ATP
14	5769.5	51.8	2261	3	Aab38111 Human ABC
15	5769.5	51.8	2261	3	Aab38114 Human ABC
16	5768.5	51.8	2261	3	Aab38117 Human ABC
17	5768.5	51.8	2261	3	Aab38115 Human ABC
18	5768.5	51.8	2261	3	Aab38109 Human ABC
19	5768.5	51.8	2261	3	Aab38082 Human ABC
20	5768.5	51.8	2261	3	Aab38112 Human ABC
21	5768.5	51.8	2261	4	Aab71749 Human ABC
22	5768.5	51.8	2261	4	Aab31361 Amino aci
23	5768.5	51.8	2261	4	Aab31365 Amino aci
24	5768.5	51.8	2261	6	Abu11899 Human ATP
25	5768.5	51.8	2261	6	Abt62034 Human ABC

26	5768.5	51.8	2261	6	ABR62033	Human ABC
27	5768.5	51.8	2261	7	ADP65173	Human ATP
28	5767.5	51.8	2261	3	AAB38116	Human ABC
29	5766.5	51.7	2261	3	AAB38113	Human ABC
30	5764.5	51.7	2261	3	AAB38110	Human ABC
31	5764.5	51.7	2261	3	AAB38105	Human ABC
32	5764.5	51.7	2261	4	AAB31362	Amino aci
33	5764.5	51.7	2261	4	AAB31366	Amino aci
34	5760.5	51.7	2261	4	AAB31363	Amino aci
35	5760.5	51.7	2261	4	AAB31367	Amino aci
36	5758	51.7	2260	3	AAB38106	Human ABC
37	5756.5	51.7	2261	3	AAB38104	Human ABC
38	5752.5	51.6	2261	5	ABB83120	Polymorph
39	5752.5	51.6	2261	5	ABB83116	Polymorph
40	5752.5	51.6	2261	5	ABB83124	Polymorph
41	5752.5	51.6	2261	5	ABB83122	Polymorph
42	5751.5	51.6	2259	3	AAB38107	Human ABC
43	5751.5	51.6	2261	4	AAM50228	Human ATP
44	5751.5	51.6	2261	4	AAB13022	Human ATP
45	5751.5	51.6	2261	4	AAU02176	Human ABC

ALIGNMENTS

RESULT 1
ABG72696
ID ABG72696 standard; protein; 2146 AA.
XX AC ABG72696;
XX AC
DT 10-MAR-2003 (first entry)
DE Human ATP-binding cassette transporter-like protein, ABCL.
KW Human; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropenia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..46
Protein	/label= Signal_peptide
	47..2146
	/label= Mature_AbCL
US2002127647-A1.	
12-SEP-2002.	
28-NOV-2001; 2001US-00995542.	
28-NOV-2000; 2000US-0253520P.	
(SHUT/) SHUTTER J.	
(ULIA/) ULIAS L.	
Shutter J, Ulia L;	
WPI; 2003-147394/14.	
N-FSDB; ABX14666.	
Novel ATP-binding cassette transporter-like polypeptides and polynucleotides useful for diagnosing, preventing, treating disorders	

PT involving immune, nervous system, thyroid, hypothalamus and impaired
FT transport of lipids.

XX Claim 13; Fig 2; 149pp; English.

XX The invention relates to an isolated murine and human ATP-binding
CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
CC encoded by the DNA insert in AtCC Deposit Nos PTA-3109, PTA-3110 or PTA-
CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
CC vectors, host cells, ABCL binding agents, a selective binding agent or
CC its fragment comprising at least one complementarity determining region
CC (CDR) with specificity for ABCL which (produced by immunising an animal
CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
CC fusion polypeptide, a device comprising a membrane suitable for
CC implantation (permeable to the protein and impermeable to materials
CC detrimental to the cells, and cells encapsulated within the membrane)
CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
CC and modulators are useful for the diagnosis and/or treatment of diseases
CC and conditions involving impaired transport of lipids, including
CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
CC involving functional and trophic disturbances of the nervous system such
CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
CC fibrosis, conditions involving multidrug resistance, conditions involving
CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
CC neutropaenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies, autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL
XX

SQ Sequence 2146 AA;

Query Match 100.0%; Score 11143; DB 6; Length 2146;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAFTQMLLLWKFMVRRRRQVQLLVLLWPLFLFVAVRSHHPLEHCHFFNKP	60
DB	1	MAFTQMLLLWKFMVRRRRQVQLLVLLWPLFLFVAVRSHHPLEHCHFFNKP	60
QY	61	LPSAGTVPWLOGLICNVNNTCFPOLTCGEPEGRLSNFNDSLVSRLLADARTVLGASAH	120
DB	61	LPSAGTVPWLOGLICNVNNTCFPOLTCGEPEGRLSNFNDSLVSRLLADARTVLGASAH	120
QY	121	TLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTSSLGALGQAQEP	180
DB	121	TLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTSSLGALGQAQEP	180
QY	181	HSLEAAEDLAQELLARLSVELRALLQRPRTSGPLELSEALCSVRGSSSTVGPSLNW	240
DB	181	HSLEAAEDLAQELLARLSVELRALLQRPRTSGPLELSEALCSVRGSSSTVGPSLNW	240
QY	241	YEASDLMELVQEPESALPOSSSPACSELIGALDSHPLSKLLWRLKPLILGKLLFAPD	300
DB	241	YEASDLMELVQEPESALPOSSSPACSELIGALDSHPLSKLLWRLKPLILGKLLFAPD	300
QY	301	TPFTTKLMAQVNRTEELTLARDVREVMELGPRIFTFMDNDSNVAMQLRLQMDGRR	360
DB	301	TPFTTKLMAQVNRTEELTLARDVREVMELGPRIFTFMDNDSNVAMQLRLQMDGRR	360
QY	361	QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVTGLGRVTECLSLDKLEAPSEAA	420
DB	361	QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLHVTGLGRVTECLSLDKLEAPSEAA	420
QY	421	LVSRALQLLAHRFWAGVVFILGPDESSDPTBHPDGLPGHVRIKIRMDIDVVRTNKIR	480
DB	421	LVSRALQLLAHRFWAGVVFILGPDESSDPTBHPDGLPGHVRIKIRMDIDVVRTNKIR	480
QY	481	DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAGLYLQOMPYPYCYDD	540

DB	481	DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAGLYLQOMPYPYCYDD	540
QY	541	VFLRVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRAMGLSRAVLWLGWFLSCL	600
DB	541	VFLRVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRAMGLSRAVLWLGWFLSCL	600
QY	601	GPFLLSAALLVILKLGDLILPYSHPGVWFLFLAAFAVATVTSQFLLSAFSFRANLAAACG	660
DB	601	GPFLLSAALLVILKLGDLILPYSHPGVWFLFLAAFAVATVTSQFLLSAFSFRANLAAACG	660
QY	661	GLAYFSLYLPYLVLCVARDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGEQAHNVG	720
DB	661	GLAYFSLYLPYLVLCVARDRLPAGGRVAASLLSPVAFGFCESLALLEEGEGEQAHNVG	720
QY	721	TRPTADVLSLAQVSGLLLDAAALYGLATWYLEAVCPQYGIPEPWNPPFRRSYWCGRPPP	780
DB	721	TRPTADVLSLAQVSGLLLDAAALYGLATWYLEAVCPQYGIPEPWNPPFRRSYWCGRPPP	780
QY	781	KSPAPCTPLDPKVLVEEAPPGLSGVSVRSLEKRPFGSPQALRGSLDLYOCHITAF	840
DB	781	KSPAPCTPLDPKVLVEEAPPGLSGVSVRSLEKRPFGSPQALRGSLDLYOCHITAF	840
QY	841	GHNGAGKTTTSLISGLFPPSGGSATILGHDVRSMAAIRPHLGVCPQYVNLFDMLTVDE	900
DB	841	GHNGAGKTTTSLISGLFPPSGGSATILGHDVRSMAAIRPHLGVCPQYVNLFDMLTVDE	900
QY	901	HMVYFGRKGLSAAVVGPEQDRLLQVGLVSKQSVQTRHLSGMQKLSVAIAFVGGSQV	960
DB	901	HMVYFGRKGLSAAVVGPEQDRLLQVGLVSKQSVQTRHLSGMQKLSVAIAFVGGSQV	960
QY	961	VILDEPTAGVDPASRRCIGIMELLIKYREGRTLIILSTHLLDEALLGDRVAVVAGRLCCG	1020
DB	961	VILDEPTAGVDPASRRCIGIMELLIKYREGRTLIILSTHLLDEALLGDRVAVVAGRLCCG	1020
QY	1021	SPFLRRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTRQKKGSGQSRVGTTPOLLA	1080
DB	1021	SPFLRRHLGSGYLLTLVKARLPLTNEKADTMEGSDVTRQKKGSGQSRVGTTPOLLA	1080
QY	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATILFRELDTRLAELRLTGYSIDTSL	1140
DB	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATILFRELDTRLAELRLTGYSIDTSL	1140
QY	1141	EIFLKVVEECAADTMEGSCGQHLCTGTAGLQVTLRLKMPPOBTALENGEPAGSAPETD	1200
DB	1141	EIFLKVVEECAADTMEGSCGQHLCTGTAGLQVTLRLKMPPOBTALENGEPAGSAPETD	1200
QY	1201	QSGSPDAVGRVQGWALTROQLQALLKRFLLARRRRRGLFAQIVLPALFVGLALVFSLIV	1260
DB	1201	QSGSPDAVGRVQGWALTROQLQALLKRFLLARRRRRGLFAQIVLPALFVGLALVFSLIV	1260
QY	1261	PPFGHYPALRLSPMTYGAQVSFFSEDA PGDPGARLLEALLQAGLEBEPVQVSHSRFSA	1320
DB	1261	PPFGHYPALRLSPMTYGAQVSFFSEDA PGDPGARLLEALLQAGLEBEPVQVSHSRFSA	1320
QY	1321	PEYPAEYAKVLASGNMTPSPSPACCSOPCARLLPDCPAAAGGPPPPQAVTSGSGVQ	1380
DB	1321	PEYPAEYAKVLASGNMTPSPSPACCSOPCARLLPDCPAAAGGPPPPQAVTSGSGVQ	1380
QY	1381	NLTGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELW	1440
DB	1381	NLTGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELW	1440
QY	1441	ALLSPLPGGALDRVLKMLTAWAHSLLDAQDSLKIWFNNKNGHWSVAFVNRASNAILLRAHP	1500
DB	1441	ALLSPLPGGALDRVLKMLTAWAHSLLDAQDSLKIWFNNKNGHWSVAFVNRASNAILLRAHP	1500
QY	1501	PGPARHAHSITTLNHPNLTKQOLSEALMASSVDVLVSVICVVFAMSFVPSFTVLIEE	1560
DB	1501	PGPARHAHSITTLNHPNLTKQOLSEALMASSVDVLVSVICVVFAMSFVPSFTVLIEE	1560
QY	1561	RVTRAKHLQLMGGSLPTLYWLGNFMDMCMNLYLPACITVLIIFLAFOQRAYVAPANLPALL	1620

Db 1561 RVTRAKHLQMGSLPTLYWLNFLNMDNCNLYLPACIVVLIPLAFQQRAYVAPANLPALL 1620
Qy 1621 LLLLYLGSITPLMYPASFFRSVPSTAYVVLTCINLFTGINGSMATFVLELPSDOKLOEV 1680
Db 1621 LLLLYLGSITPLMYPASFFRSVPSTAYVVLTCINLFTGINGSMATFVLELPSDOKLOEV 1680
Qy 1681 SRLKQVFLIPPHFCGLGRGLIDMVRNQAMADAFERLGRQFQSPLRWEVVGKLLAMVIQ 1740
Db 1681 SRLKQVFLIPPHFCGLGRGLIDMVRNQAMADAFERLGRQFQSPLRWEVVGKLLAMVIQ 1740
Qy 1741 GFLFLLFTLLQHRSQLPQPRVRSPLPGLGEDEDAERERERVOGATQGDVLRNLTK 1800
Db 1741 GFLFLLFTLLQHRSQLPQPRVRSPLPGLGEDEDAERERERVOGATQGDVLRNLTK 1800
Qy 1801 VYRGQMPAVDRCLGIPPGCEFGLLGVNGACKTSTFRVWTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VYRGQMPAVDRCLGIPPGCEFGLLGVNGACKTSTFRVWTGDTLASRGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1920
Db 1861 EPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1920
Qy 1921 RPAGTYSNGNKKLATALALVGDPAVFLDEPTTGMDSARRFLWNSLLAVVREGRSVWL 1980
Db 1921 RPAGTYSNGNKKLATALALVGDPAVFLDEPTTGMDSARRFLWNSLLAVVREGRSVWL 1980
Qy 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Db 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Qy 2041 EPPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
Db 2041 EPPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
Qy 2101 SKDQKDEDTBEQKEAGVGDPAVGLQHPKRVSQFLDDPSTAEVTL 2146
Db 2101 SKDQKDEDTBEQKEAGVGDPAVGLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 2
AAU04483

ID AAU04483 standard; protein; 2146 AA.

XX AC AAU04483;

XX DT 26-SEP-2001 (first entry)

XX DE Human PD-ATP-binding cassette (PD-ABC) protein form #1.

XX KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.

XX OS Homo sapiens.

XX PN WO200153490-A1.

XX PD 26-JUL-2001.

XX PF 23-JAN-2001; 2001WO-US002191.

XX PR 24-JAN-2000; 2000US-0177889P.

XX PR 30-JUN-2000; 2000US-0215405P.

XX PA (WARN) WARNER LAMBERT CO.

XX PI Johns MA, Tafuri SR, Wang M;

XX DR WPI; 2001-442259/47.

DR N-PSDB; AAS08706.

XX PT New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
PT of dyslipidaemia, epilepsy and diseases related to abnormal calcium flux.

XX PS Claim 10; Page 54-62; 77pp; English.

XX CC The sequence represents human PD-ATP-binding cassette (PD-ABC) protein
form 1. PD-ABC maps to chromosome 19p13.3 and is expressed in various
tissues including spleen, thymus, peripheral blood leukocytes, bone
marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to
diagnose and treat cardiovascular disorders, inflammatory disorders,
dyslipidaemia, epilepsy, diseases related to abnormal calcium flux,
coronary artery disease, Tangier's disease, familial high-density
lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
insulin resistance, obesity, alcoholism, retinal degeneration,
hypertension and vascular disease. The sequences are also used in drug
screening assays

XX SQ Sequence 2146 AA;

Query Match 99.9%; Score 11130; DB 4; Length 2146;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2144; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAFWTQMLLWKMFYRRROPVOLLVELLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60

Db 1 MAFWTQMLLWKMFYRRROPVOLLVELLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60

Qy 61 LPSAGTVPMWLGGLICNVNNTCFPQITPGEEPRGLSNFNDLSVRLADARTVLGGASAH 120

Db 61 LPSAGTVPMWLGGLICNVNNTCFPQITPGEEPRGLSNFNDLSVRLADARTVLGGASAH 120

Qy 121 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMDLVAELLTSLRTESLGLALGOAQL 180

Db 121 TLAGLGLKIATLRAARSTAOPTKQSPLEPPMDLVAELLTSLRTESLGLALGOAQL 180

Qy 181 HSLEAEADLAQELLALRSVELRALLORPGTSGPLELLSEALCSVRGPSTVGPSSLNW 240

Db 181 HSLEAEADLAQELLALRSVELRALLORPGTSGPLELLSEALCSVRGPSTVGPSSLNW 240

Qy 241 YEASDLMELVQEPESALPDSSLPACSELIGALDSHPLSRLMRLLKPLILGLLFPD 300

Db 241 YEASDLMELVQEPESALPDSSLPACSELIGALDSHPLSRLMRLLKPLILGLLFPD 300

Qy 301 TPFTKRLMAQVNRTEFELTLRLDVRVWEMLGPRITFTPMNDSSNVAMLQRLLOMDEGR 360

Db 301 TPFTKRLMAQVNRTEFELTLRLDVRVWEMLGPRITFTPMNDSSNVAMLQRLLOMDEGR 360

Qy 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTGRVTECLSLDKLEAAPSSAA 420

Db 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTGRVTECLSLDKLEAAPSSAA 420

Qy 421 LVSRALQLLAEHRFWAGVVFGLGEDSDPTPEHPTDLPQPGHVRKIRMDIDVVTNRKIR 480

Db 421 LVSRALQLLAEHRFWAGVVFGLGEDSDPTPEHPTDLPQPGHVRKIRMDIDVVTNRKIR 480

Qy 481 DRFWDGCPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYPCYVDD 540

Db 481 DRFWDGCPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYPCYVDD 540

Qy 541 VFLRVLSRSLPFLTLTAWIYSVTLTVKAVVREKETRLDTRAMGLSAVLWGLWFLSCL 600

Db 541 VFLRVLSRSLPFLTLTAWIYSVTLTVKAVVREKETRLDTRAMGLSAVLWGLWFLSCL 600

Qy 601 GPFLLSAALLVLKLGDIILFYSHPGVVFLEAFAVATVTSFLLSAFFSRANLAAACG 660

Db 601 GPFLLSAALLVLKLGDIILFYSHPGVVFLEAFAVATVTSFLLSAFFSRANLAAACG 660

Qy 661 GLAYFSLYLPVYLCVAVWRDLRPLAGRVAASLLSPVAFGCGESIALLEEQEGAGHNVG 720

Db 661 GLAYFSLYLPVYLCVAVWRDLRPLAGRVAASLLSPVAFGCGESIALLEEQEGAGHNVG 720

QY 721 TRPTADVSLAQSGILLDDAALYGLATWYLEAVCPQYGIPEPWNPFRRSVCWGRPP 780
DB 721 TRPTADVSLAQSGILLDDAALYGLATWYLEAVCPQYGIPEPWNPFRRSVCWGRPP 780
QY 781 KSPAPCPTLPDPKVLVEEAPPGLSPGVSVRSLEKRRFPQALRGISLDFYQGHITAF 840
DB 781 KSPAPCPTLPDPKVLVEEAPPGLSPGVSVRSLEKRRFPQALRGISLDFYQGHITAF 840
QY 841 CHNGAGKTTTILSILSGLFPSPGGSFAFLGHDRSSMAAIRPHLGVCQYNNVLFDMLTVD 900
DB 841 CHNGAGKTTTILSILSGLFPSPGGSFAFLGHDRSSMAAIRPHLGVCQYNNVLFDMLTVD 900
QY 901 HVMFYGRLKGLSAVGPEDRLQDVLGVSQSVQTRHLSGGMORKLSVAIAFVGSQV 960
DB 901 HVMFYGRLKGLSAVGPEDRLQDVLGVSQSVQTRHLSGGMORKLSVAIAFVGSQV 960
QY 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLLIILSTHLDDEALLGDRVAVVAGRLCCCG 1020
DB 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLLIILSTHLDDEALLGDRVAVVAGRLCCCG 1020
QY 1021 SPLFLRRHLGSGYYLLTVKARLPLTTNEKADTMEGSDVTROEKKNGSQSRRVGT 1080
DB 1021 SPLFLRRHLGSGYYLLTVKARLPLTTNEKADTMEGSDVTROEKKNGSQSRRVGT 1080
QY 1081 LVQHWPFGARLVEELPHELVLVLPYTCAGHDSFATLRELDTLAEALRTGYISDTSLE 1140
DB 1081 LVQHWPFGARLVEELPHELVLVLPYTCAGHDSFATLRELDTLAEALRTGYISDTSLE 1140
QY 1141 EIFLKVVVEECAAADTMDGSCGQHLCTGAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
DB 1141 EIFLKVVVEECAAADTMDGSCGQHLCTGAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARRSRGLFAQIVLPALFVGLALVFSLI 1260
DB 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARRSRGLFAQIVLPALFVGLALVFSLI 1260
QY 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGPRARLLEALLQEALEPPVQSHSRFSA 1320
DB 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGPRARLLEALLQEALEPPVQSHSRFSA 1320
QY 1321 PEVPAEVAKLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQ 1380
DB 1321 PEVPAEVAKLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQ 1380
QY 1381 NLTCGRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSGLGRDPLPSQELGRSVEELW 1440
DB 1381 NLTCGRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSGLGRDPLPSQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTAWAHSIDQAQDSLKIWFNNKGWHSWVAFVNRASNAILRAHLP 1500
DB 1441 ALLSPLPGGALDRVLKNTAWAHSIDQAQDSLKIWFNNKGWHSWVAFVNRASNAILRAHLP 1500
QY 1501 PGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVCVFAMSFPVPSFTLVLEE 1560
DB 1501 PGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVCVFAMSFPVPSFTLVLEE 1560
QY 1561 RVTRAKHLQMGGLSPTLYLWGNFLWDMCNVLPACIVLIFLAQOQRAYVAPANLPAL 1620
DB 1561 RVTRAKHLQMGGLSPTLYLWGNFLWDMCNVLPACIVLIFLAQOQRAYVAPANLPAL 1620
QY 1621 LLLLYGWSITPLMPYASFPFSPSTAYVVLVTCINLFIGINGSMATFVLELPSDKLOEV 1680
DB 1621 LLLLYGWSITPLMPYASFPFSPSTAYVVLVTCINLFIGINGSMATFVLELPSDKLOEV 1680
QY 1681 SRILKQVFLIPFHCCLRGILDMVRNOAMADAPERLGDQFQSPPLRWEVVGKILLAMVIO 1740
DB 1681 SRILKQVFLIPFHCCLRGILDMVRNOAMADAPERLGDQFQSPPLRWEVVGKILLAMVIO 1740
QY 1741 GPFLEFLLQLQHSQQLLPQPRVRSPLFLGGEEDVARERVVQATQGDVLVRLNLT 1800
DB 1741 GPFLEFLLQLQHSQQLLPQPRVRSPLFLGGEEDVARERVVQATQGDVLVRLNLT 1800
QY 1801 VYRQORMPAVDRCLGIPPGECFGLLVNGVAGKSTSTFRMTGDTLASRGEAVLAGHSVAR 1860

DB 1801 VYRQORMPAVDRCLGIPPGECFGLLVNGVAGKSTSTFRMTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLHSGVYCPQSDAIPELLTGRHEHLELLARLRGVPEAOVAQTAGSLARLGLSWYAD 1920
DB 1861 EPSAAHLHSGVYCPQSDAIPELLTGRHEHLELLARLRGVPEAOVAQTAGSLARLGLSWYAD 1920
QY 1921 RPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
DB 1921 RPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
QY 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTRVPAARSQPAAAFVAA 2040
DB 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTRVPAARSQPAAAFVAA 2040
QY 2041 EFPGSBELRAHGRLRFQLPFGRCALARVFGELAVHGAHGVDFSVSQTMLVEEVFLYF 2100
DB 2041 EFPGSBELRAHGRLRFQLPFGRCALARVFGELAVHGAHGVDFSVSQTMLVEEVFLYF 2100
QY 2101 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL 2146
DB 2101 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL 2146
RESULT 3
AA014210
ID AA014210 standard; protein; 2180 AA.
XX AA014210;
XX AC AA014210;
XX DT 03-MAY-2002 (first entry)
XX DE Human transporter and ion channel TRICH-27.
XX KW Human; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder;
KW cell proliferative disorder; neuroprotective; neotropic;
KW cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
KW gene therapy.
XX OS Homo sapiens.
XX PN WO200204520-A2.
XX PD 17-JAN-2002.
XX PF 05-JUL-2001; 2001WO-US021448.
XX PR 07-JUL-2000; 2000US-0216547P.
PR 14-JUL-2000; 2000US-0218232P.
PR 21-JUL-2000; 2000US-0220112P.
PR 28-JUL-2000; 2000US-0221839P.
XX (INCY-) INCYTE GENOMICS INC.
XX RAumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J;
PI Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Wallia NK;
PI Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
PI Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;
PI Kearney L, Thangavelu K, Das D, Policky JL;
XX WPI; 2002-205969/26.
DR N-PSDB; AAL44693.
XX New human transporters and ion channel polypeptides for diagnosing,
PT treating or preventing transport, neurological, muscle, immunological and
PT cell proliferative disorders.
XX Claim 1; Page 188-192; 230pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of human transporter and ion channel proteins, designated TRICH-1-

CC TRICH-32. The sequences can be used in the treatment of transport, CC neurological, muscle, immunological and cell proliferative disorders. The CC present sequence is a protein of the invention

XX Sequence 2180 AA;

QY	1	MAFWTQMLLLWKNFMYRRQPVQLLVLLWPLFLFFILVAVVRHSHPPLEHHECHFPNKP	60
DB	35	MAFWTQMLLLWKNFMYRRQPVQLLVLLWPLFLFFILVAVVRHSHPPLEHHECHFPNKP	94
QY	61	LPSAGTVPMLOGLICNVNNTCPQPTPGEPGRLSNFNDSLVSRLLADARTVVGASAH	120
DB	95	LPSAGTVPMLOGLICNVNNTCPQPTPGEPGRLSNFNDSLVSRLLADARTVVGASAH	154
QY	121	TLAIGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQORPL	180
DB	155	TLAIGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQORPL	214
QY	181	HSLLEAAEDLAQELALRSLEIRALLQRPRTSGPLELISEALCSVRGSPSTVGPSLNW	240
DB	215	HSLLEAAEDLAQELALRSLEIRALLQRPRTSGPLELISEALCSVRGSPSTVGPSLNW	274
QY	241	YEASDLMELVGPESALPDSSLPACSELIGALDHPRLSRLWRLKPLILGKLLFAPD	300
DB	275	YEASDLMELVGPESALPDSSLPACSELIGALDHPRLSRLWRLKPLILGKLLFAPD	334
QY	301	TPFTTKLMAQVNRTEELTLRDREVWEMGLPRIFTFMNDSSNVAMLQRLLOMDEGRR	360
DB	335	TPFTTKLMAQVNRTEELTLRDREVWEMGLPRIFTFMNDSSNVAMLQRLLOMDEGRR	394
QY	361	QRPGRDHMEALRFLDPGSCGYQWODAHADVHLVGTLRVTSCLSDKLBAAPSEAA	420
DB	395	QRPGRDHMEALRFLDPGSCGYQWODAHADVHLVGTLRVTSCLSDKLBAAPSEAA	454
QY	421	LVSRALQLLAEHRFWAGVVFVLPEDSSDPTHTPDGLGPHVRIKIRMDIDVVTNKIR	480
DB	455	LVSRALQLLAEHRFWAGVVFVLPEDSSDPTHTPDGLGPHVRIKIRMDIDVVTNKIR	514
QY	481	DRFWDGPAADPLTLRLVYWGFFVYLQDLVERAAVRLSGANPRAGLYLQMPYPCYVDD	540
DB	515	DRFWDGPAADPLTLRLVYWGFFVYLQDLVERAAVRLSGANPRAGLYLQMPYPCYVDD	574
QY	541	VPLRLVRSRLPLFLTLAWIYSVTLTVKAVVRKETRLRDTWRAMGLSRAVLWGLWFLSCL	600
DB	575	VPLRLVRSRLPLFLTLAWIYSVTLTVKAVVRKETRLRDTWRAMGLSRAVLWGLWFLSCL	634
QY	601	GPFLLSAALLVLAVLKGDLPLVSHPGVVFLFLAAFAVATVTSQFLLSAFFGRANLAAACG	660
DB	635	GPFLLSAALLVLAVLKGDLPLVSHPGVVFLFLAAFAVATVTSQFLLSAFFGRANLAAACG	694
QY	661	GLAYFSLXPLPYVLVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEQEGEAQHNVG	720
DB	695	GLAYFSLXPLPYVLVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEQEGEAQHNVG	754
QY	721	TRPTADVFLAQVSGILLDLDAALYGLATWYLEAVCPQGYIPEPNWFFRFRSYWCGPRPP	780
DB	755	TRPTADVFLAQVSGILLDLDAALYGLATWYLEAVCPQGYIPEPNWFFRFRSYWCGPRPP	814
QY	781	KSPAPCPTPLDPKVLVEAPPLGSLPGSVRSLEKEFRGSPQALRGLSLDFVQGHITAFI	840
DB	815	KSPAPCPTPLDPKVLVEAPPLGSLPGSVRSLEKEFRGSPQALRGLSLDFVQGHITAFI	874
QY	841	GHNGAGKTTTILSGLPFPSPGSAFILGHVDVSRSSMAAIRPHLGVCPQYNVFLFDMLTVD	900
DB	875	GHNGAGKTTTILSGLPFPSPGSAFILGHVDVSRSSMAAIRPHLGVCPQYNVFLFDMLTVD	934
QY	901	HWVFYGRUKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQRKLSVAIAFVGGSQV	960
DB	935	HWVFYGRUKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQRKLSVAIAFVGGSQV	994

QY	961	VILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDDEALLGDRVAVVAGGRCCCG	1020
DB	995	VILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDDEALLGDRVAVVAGGRCCCG	1054
QY	1021	SPLFRRHLGGYYITLVKARLPLTTNEKADTMEGSDVTROEKXNGSGSRVGTPLL	1080
DB	1055	SPLFRRHLGGYYITLVKARLPLTTNEKADTMEGSDVTROEKXNGSGSRVGTPLL	1114
QY	1081	LVQHWVPCARLVEELPHELVLVLPYTGADGSPATLFRDLTRLAELRLTGVGISDTSLE	1140
DB	1115	LVQHWVPCARLVEELPHELVLVLPYTGADGSPATLFRDLTRLAELRLTGVGISDTSLE	1174
QY	1141	EIFLVKVECAADTMEGSCGQHLCTCTGIAGLDVTLRLKMPPOETALENGEPAGAPETD	1200
DB	1175	EIFLVKVECAADTMEGSCGQHLCTCTGIAGLDVTLRLKMPPOETALENGEPAGAPETD	1234
QY	1201	QSGSDPAVGRVQGWALTQQLQALLKFLIARRSRRLFAQIVLPALFVGLALVFSLTIV	1260
DB	1235	QSGSDPAVGRVQGWALTQQLQALLKFLIARRSRRLFAQIVLPALFVGLALVFSLTIV	1294
QY	1261	PPRGHPALRLSPTMYGAOVFFSSEDAPGDPGRARLLLEALLOEAGLEBPPVQSHSRFSA	1320
DB	1295	PPRGHPALRLSPTMYGAOVFFSSEDAPGDPGRARLLLEALLOEAGLEBPPVQSHSRFSA	1354
QY	1321	PEVPAEVAKVLASGNWTPESPSPACQSPQAGARRLLPDCPAAAGGPPPPQAVTGSVEVQ	1380
DB	1355	PEVPAEVAKVLASGNWTPESPSPACQSPQAGARRLLPDCPAAAGGPPPPQAVTGSVEVQ	1414
QY	1381	NLTGNLSDFLVKTYPRLVROGLTKKWVNEVYRGSFSLGGRDPLGSGQELGRSVEELW	1440
DB	1415	NLTGNLSDFLVKTYPRLVROGLTKKWVNEVYRGSFSLGGRDPLGSGQELGRSVEELW	1474
QY	1441	ALLSPLPGALDRVLKNTAWAHSDDAODSLKIFNNKNGWSMAFVNRASNAIIRAHLP	1500
DB	1475	ALLSPLPGALDRVLKNTAWAHSDDAODSLKIFNNKNGWSMAFVNRASNAIIRAHLP	1534
QY	1501	PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSI CVVPAMSFPVPSFTLVLTBE	1560
DB	1535	PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSI CVVPAMSFPVPSFTLVLTBE	1594
QY	1561	RVTRAKHLQMLGGLSPTLYLWLNFLWDMCNLYLPACI VVLI FLAFQRAYVAPANLPALL	1620
DB	1595	RVTRAKHLQMLGGLSPTLYLWLNFLWDMCNLYLPACI VVLI FLAFQRAYVAPANLPALL	1654
QY	1621	LLLLLYGWSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELPSDQKLOEV	1680
DB	1655	LLLLLYGWSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELPSDQKLOEV	1714
QY	1681	SRIILKQVFLIFPHFCLGRGLDMVRNOAMADAFERLGRDQFQSPURWEVGNLWAMVQ	1740
DB	1715	SRIILKQVFLIFPHFCLGRGLDMVRNOAMADAFERLGRDQFQSPURWEVGNLWAMVQ	1774
QY	1741	GPFLFLLTLLQHSRQLLPQPRVRSPLLGGBEDVARERERVQATGQGDVILNRLNLT	1800
DB	1775	GPFLFLLTLLQHSRQLLPQPRVRSPLLGGBEDVARERERVQATGQGDVILNRLNLT	1834
QY	1801	VYRGQMPADVRLCLIGIPPGECFGLLVNGAGKTSTFRMTGDTTLASGEAVLAGHSVAR	1860
DB	1835	VYRGQMPADVRLCLIGIPPGECFGLLVNGAGKTSTFRMTGDTTLASGEAVLAGHSVAR	1894
QY	1861	EPSSAHLNMGYCPQSDAI FELLTGREHLELLARLGRVPEAOVACTAGSLARLGLSWAD	1920
DB	1895	EPSSAHLNMGYCPQSDAI FELLTGREHLELLARLGRVPEAOVACTAGSLARLGLSWAD	1954
QY	1921	RPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARRFLWNSLLAVVREGSVML	1980
DB	1955	RPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARRFLWNSLLAVVREGSVML	2014
QY	1981	TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAFVAA	2040
DB	2015	TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAFVAA	2074

QY 2041 EFGSELREAHGGRRLRQLPPGRCALARVFGELAVHGAHGVDFSVSTMLEEFLYF 2100
 DB 2075 EFGAEIREAHGGRRLRQLPPGRCALARVFGELAVHGAHGVDFSVSTMLEEFLYF 2134
 QY 2101 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSGFLDDPSTAETVL 2146
 DB 2135 SKDQKDEDETEQKEAGVGVDPAPGLQHPKRVSGFLDDPSTAETVL 2180

RESULT 4
 ABP52096
 ID ABP52096 standard; protein; 2146 AA.
 XX
 AC ABP52096;
 XX
 DT 10-OCT-2002 (first entry)
 XX
 DE Homo sapiens ABC transporter ABCA7 protein SEQ ID NO:48.
 XX
 KW ATP-binding cassette transporter; ABC transporter; modulation; D loop;
 KW cancer; bacterial infection; fungal infection; protozoal infection;
 KW antibacterial; fungicide; protozoacide.
 XX
 OS Homo sapiens.
 XX
 PN EPI217066-A1.
 XX
 PD 26-JUN-2002.
 XX
 PF 21-DEC-2000; 2000EP-00870316.
 XX
 PR 21-DEC-2000; 2000EP-00870316.
 XX
 PA (UYGE-) UNIV GENT.
 XX
 XX WPI; 2002-550404/59.
 XX
 PT Modulating activity of ATP-binding cassette (ABC) transporters by
 PT influencing dimerization of nucleotide binding domains through use of D
 PT loop sequence of an ABC transporter, or its antisense peptide or peptide
 PT mimetic.
 XX
 PS Disclosure; Fig 3; 290pp; English.
 XX
 CC The present invention describes a method (M1) for modulating the activity
 CC of ATP-binding cassette (ABC) transporters by influencing the
 CC dimerization of the nucleotide binding domains comprising using: (a) a
 CC polypeptide (polyp) consisting of 5-50 amino acids comprising the D loop
 CC sequence of an ABC transporter (ABP52049 to ABP52091); (b) a polyp
 CC consisting of the D loop sequence of an ABC transporter; (c) a peptide
 CC mimetic or antisense peptide of (a) or (b). ABC transporters have
 CC antibacterial, fungicide and protozoacide activities. (M1) is useful for
 CC selectively modulating the activity of ABC transporters belonging to the
 CC group of multidrug transporter/P-glycoproteins. Bacterial, fungal or
 CC protozoal ABC transporters are involved in the infection of a mammal or
 CC in the induction of resistance to antibiotics or drugs in a mammal. (M1)
 CC is useful for preventing, treating or alleviating diseases associated
 CC with functionality of an ABC transporter. ABP52092 to ABP52140 represent
 CC ABC transporter proteins given in the exemplification of the present
 CC invention

QY 1 MAFWTQMLLWKNNFRRRROPVQLLVLLWPLFLFLLVAVRSHHPLEHCHFNKP 60
 DB 1 MAFWTQMLLWKNNFRRRROPVQLLVLLWPLFLFLLVAVRSHHPLEHCHFNKP 60
 QY 61 LPSAGTVPWLQGLICNVNNTCFPOLTPGEEPRGLSNFNDLSVSRLLADARTVLGASAH 120

DB 61 LPSAGTVPWLQGLICNVNNTCFPOLTPGEEPRGLSNFNDLSVSRLLADARTVLGASAH 120
 QY 121 TLAGLCKLTATLRAASTAQPOPTKOSPLEPPMLDVAELLTSLRTTESLGLAQOEPL 180
 DB 121 TLAGLCKLTATLRAASTAQPOPTKOSPLEPPMLDVAELLTSLRTTESLGLAQOEPL 180
 QY 181 HSLEAAEDLAQELLALRSLSVELRALLQRPRTSGPLELSEALCSVRGSPSSVGSPLNW 240
 DB 181 HSLEAAEDLAQELLALRSLSVELRALLQRPRTSGPLELSEALCSVRGSPSSVGSPLNW 240
 QY 241 YEASDLMELVGPESALPDSSLPACSELIGALDSHPLSRLLRRLKPLILKLLFAPD 300
 DB 241 YEASDLMELVGPESALPDSSLPACSELIGALDSHPLSRLLRRLKPLILKLLFAPD 300
 QY 301 TPTFKLMAQVNTPEELTLRDVREVMELGPRIETFMNDSSNVAMQLQLOQDEGR 360
 DB 301 TPTFKLMAQVNTPEELTLRDVREVMELGPRIETFMNDSSNVAMQLQLOQDEGR 360
 QY 361 QRPFGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
 DB 361 QRPFGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
 QY 421 LVSRAQLLAEHRFWAGVFLGPEDSDPTEHPTDLPGLPGHVRIRKIMDIDVVRTNKIR 480
 DB 421 LVSRAQLLAEHRFWAGVFLGPEDSDPTEHPTDLPGLPGHVRIRKIMDIDVVRTNKIR 480
 QY 481 DRFWDGPAADPLTDLRYVWGGVYLODLVERAAVRVLSGANPRAGLYLOOMPYPVCVDD 540
 DB 481 DRFWDGPAADPLTDLRYVWGGVYLODLVERAAVRVLSGANPRAGLYLOOMPYPVCVDD 540
 QY 541 VFLRVLSRLPLFLTLAWIYSVTLTVKAHVREKETRLDTRAMGLSRAVLWGLFSLCL 600
 DB 541 VFLRVLSRLPLFLTLAWIYSVTLTVKAHVREKETRLDTRAMGLSRAVLWGLFSLCL 600
 QY 601 GPFLLSAALLVVLKLDILPYSHPGVFLFLAAFAVATVTSFLLSAFSSRANLAAACG 660
 DB 601 GPFLLSAALLVVLKLDILPYSHPGVFLFLAAFAVATVTSFLLSAFSSRANLAAACG 660
 QY 661 GLAYFSLYLPYVLCVAVRDELPAAGRAVASLLSPVAFGFCESLALLEEGEGQAQHNVG 720
 DB 661 GLAYFSLYLPYVLCVAVRDELPAAGRAVASLLSPVAFGFCESLALLEEGEGQAQHNVG 720
 QY 721 TRPTADVLSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYIPEPMPNPPFRSSWCGPRPP 780
 DB 721 TRPTADVLSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYIPEPMPNPPFRSSWCGPRPP 780
 QY 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLDFYQGHITAF 840
 DB 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLDFYQGHITAF 840
 QY 841 GHNGAGKTTTSLTSLGSLFPPSGGSAFILGHDVRSMAAIPHLGVCPOYNNVLDMLTVDE 900
 DB 841 GHNGAGKTTTSLTSLGSLFPPSGGSAFILGHDVRSMAAIPHLGVCPOYNNVLDMLTVDE 900
 QY 901 HVMFYGRKLGSLAAVVGPEODRLLODVGLVSKSVOTRHLSGWQKLSVAIAFVGSQV 960
 DB 901 HVMFYGRKLGSLAAVVGPEODRLLODVGLVSKSVOTRHLSGWQKLSVAIAFVGSQV 960
 QY 961 VIIDEPTAGVDPAARRGIMWELLKRYEGRTLLIISTHLDDEALLGDRVAVVAGRLCCCG 1020
 DB 961 VIIDEPTAGVDPAARRGIMWELLKRYEGRTLLIISTHLDDEALLGDRVAVVAGRLCCCG 1020
 QY 1021 SPLFLRRHLGSGYLLTVLKARLPLTTNEKADTMEGSDVTRQEKNGSQSRVGTPOLLA 1080
 DB 1021 SPLFLRRHLGSGYLLTVLKARLPLTTNEKADTMEGSDVTRQEKNGSQSRVGTPOLLA 1080
 QY 1081 LVQHWPGARLVSELPHLVLVLPYGAHDSFATLFLRELDTRLAEURLGYGISTDSLE 1140
 DB 1081 LVQHWPGARLVSELPHLVLVLPYGAHDSFATLFLRELDTRLAEURLGYGISTDSLE 1140
 QY 1141 EIFLKVVEECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
 DB 1141 EIFLKVVEECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200

```
QY 1201 QSGGPDVGRVQGWALTRQOQALLKRRFLARRSRRLGPAQIVLPALFVGLALVPSLIV 1260
Db 1201 QSGGPDVGRVQGWALTRQOQALLKRRFLARRSRRLGPAQIVLPALFVGLALVPSLIV 1260
QY 1261 PPGHYPALRLSPMTYGAQVSPFSEDAPDGPGRARLEALLQAGLEEPVQSHSHRFA 1320
Db 1261 PPGHYPALRLSPMTYGAQVSPFSEDAPDGPGRARLEALLQAGLEEPVQSHSHRFA 1320
QY 1321 PEVPAEAVKVLASGNWTPSPSPACOCOPGARRLLPDCPAAAGGPPPPQAVTGSQEVQ 1380
Db 1321 PEVPAEAVKVLASGNWTPSPSPACOCOPGARRLLPDCPAAAGGPPPPQAVTGSQEVQ 1380
QY 1381 NUTGRNLSDFLVKTYPRLVKTKKWNVEVRYGSLGGRDPGLPSQELGRSVEELW 1440
Db 1381 NUTGRNLSDFLVKTYPRLVKTKKWNVEVRYGSLGGRDPGLPSQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTAWAHSLSDAQSLKIFWNNKGWHSWVAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSLSDAQSLKIFWNNKGWHSWVAFVNRASNAILRAHLP 1500
QY 1501 PGPARRHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLE 1560
Db 1501 PGPARRHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLE 1560
QY 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNLYLPACIVVLIIFLAFOORAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNLYLPACIVVLIIFLAFOORAYVAPANLPALL 1620
QY 1621 LLLLYLWGSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLQV 1680
Db 1621 LLLLYLWGSITPLMPYSPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLQV 1680
QY 1681 SRLKQVFLIFPHFCIGRLIDMVRNOAMADAFERLGFQFOSPLRWEVVGKLLAMVQ 1740
Db 1681 SRLKQVFLIFPHFCIGRLIDMVRNOAMADAFERLGFQFOSPLRWEVVGKLLAMVQ 1740
QY 1741 GFLFLFTLLQHRSQLPQVRSLPLGGEDEDEVARERERVQATGQDVLVRLNLT 1800
Db 1741 GFLFLFTLLQHRSQLPQVRSLPLGGEDEDEVARERERVQATGQDVLVRLNLT 1800
QY 1801 VTRGQMPADVRLCLGIPGECFGLLVGNVAGKTTSTFRMTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VTRGQMPADVRLCLGIPGECFGLLVGNVAGKTTSTFRMTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVATAGSLARLGSWAD 1920
Db 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVATAGSLARLGSWAD 1920
QY 1921 RPAGTYSNGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLWNSLLAVVREGSVML 1980
Db 1921 RPAGTYSNGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLWNSLLAVVREGSVML 1980
QY 1981 TSHSMECEALCSRLAIVMNGFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Db 1981 TSHSMECEALCSRLAIVMNGFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
QY 2041 EPPGSELREAHGGRRLFQLPPGRCALARVFGELAVHGAHEGVDFSVQTMLEEVFLYF 2100
Db 2041 EPPGSELREAHGGRRLFQLPPGRCALARVFGELAVHGAHEGVDFSVQTMLEEVFLYF 2100
QY 2101 SKDQKDEDETEREKAGVGDVPAPGLQHPKRVYSQFLDDPSTAEVT 2146
Db 2101 SKDQKDEDETEREKAGVGDVPAPGLQHPKRVYSQFLDDPSTAEVT 2146
```

RESULT 5

ABU08466

ID ABU08466 standard; protein; 2146 AA.

XX AC ABU08466;

XX DT 18-JUN-2003 (first entry)

```
XX Human ABCA-SSN protein.
DE Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory; ABCA-SSN;
KW antiarteriosclerotic.
XX Homo sapiens.
XX WO2003010315-A1.
XX 06-FEB-2003.
XX 24-JUL-2002; 2002WO-JP007487.
XX 25-JUL-2001; 2001JP-00224176.
XX 06-DEC-2001; 2001JP-00372530.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX (KAZU-) KAZUSA DNA RES INST FOUND.
XX Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX Claim 18; Page 164-174; 183pp; Japanese.
XX The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA-SSN protein
XX Sequence 2146 AA;
Query Match 99.9%; Score 11128; DB 6; Length 2146;
Best Local Similarity 99.9%; Pred No. 0;
Matches 2144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAFWTQMLLLWKPMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MAFWTQMLLLWKPMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
QY 61 LPSAGTVPWLOGLICNVNNTCFPQLTGPEEPGRLSNFNDSLSVRLADARTVLGASAH 120
Db 61 LPSAGTVPWLOGLICNVNNTCFPQLTGPEEPGRLSNFNDSLSVRLADARTVLGASAH 120
QY 121 TLGAGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQAEPL 180
Db 121 TLGAGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQAEPL 180
QY 181 HSLLEAAEDLAQELLARSLVELRALLQRPRTSGPFLLESLALCSVRGSPSTVGP 240
Db 181 HSLLEAAEDLAQELLARSLVELRALLQRPRTSGPFLLESLALCSVRGSPSTVGP 240
QY 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRLLKPLIIGKLLFAPD 300
Db 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRLLKPLIIGKLLFAPD 300
QY 301 TPFTKLMQAVNRTFEELTLDRDREVWEMLGPRIFTFMNDSNVAMLQRLQMDQGR 360
Db 301 TPFTKLMQAVNRTFEELTLDRDREVWEMLGPRIFTFMNDSNVAMLQRLQMDQGR 360
QY 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHVGLVGTGLGRVTCLSLDKLEAAPSEAA 420
```

Db 361 QPRGGRDHMEALRSFLDPGSGGYSQDAHADVGHVLTGLGRVTECLSLDKLEAAPSEAA 420
Qy 421 LVSRALQOLLAHRFWAGVVFGLPDESDPTHEPDPDLGPGHVRIRKIMDIDVTRTNKIR 480
Db 421 LVSRALQOLLAHRFWAGVVFGLPDESDPTHEPDPDLGPGHVRIRKIMDIDVTRTNKIR 480
Qy 481 DRFWDGPAADPLTDLRVVGGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPVPCVYDD 540
Db 481 DRFWDGPAADPLTDLRVVGGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPVPCVYDD 540
Qy 541 VFLRVLSRSLPFLTLTAWISYVTLTVKAVVREKETRLRDTWRAMGLSRAVLMLGWFLSCL 600
Db 541 VFLRVLSRSLPFLTLTAWISYVTLTVKAVVREKETRLRDTWRAMGLSRAVLMLGWFLSCL 600
Qy 601 GPFLLSALILVLKLGDIILPYSHPGVVFELFLAAFAVATVTSQFLLSAFFSRANLAAACG 660
Db 601 GPFLLSALILVLKLGDIILPYSHPGVVFELFLAAFAVATVTSQFLLSAFFSRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVAMRDLRPAAGRVAASLLSPVAFGFCESLALLEEGEGQAHNVG 720
Db 661 GLAYFSLYLPVLCVAMRDLRPAAGRVAASLLSPVAFGFCESLALLEEGEGQAHNVG 720
Qy 721 TRPTADVFSLAQVSGLLLDAAALYGLATWILEAVCPQYQYGIPEPWNPFPRRSYWCGRPP 780
Db 721 TRPTADVFSLAQVSGLLLDAAALYGLATWILEAVCPQYQYGIPEPWNPFPRRSYWCGRPP 780
Qy 781 KSPAPCTPLDPKVLVEAPPGSLSPGVSRSLERKRPFGSPQALRGISLDFYQGHITAF 840
Db 781 KSPAPCTPLDPKVLVEAPPGSLSPGVSRSLERKRPFGSPQALRGISLDFYQGHITAF 840
Qy 841 GHNGAGTKTLLSILSGLFPFGSGSAFLLGHDRSSMAAIRPHLGVCPOYNVLFPLMTVDE 900
Db 841 GHNGAGTKTLLSILSGLFPFGSGSAFLLGHDRSSMAAIRPHLGVCPOYNVLFPLMTVDE 900
Qy 901 HWFYGLKGLSAVVGPEODRLQDVLVSKQSVQTRHLSGGNQKLSVAIAFVGGSQV 960
Db 901 HWFYGLKGLSAVVGPEODRLQDVLVSKQSVQTRHLSGGNQKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLIILSTHLDABELLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIIWELLKYREGRTLIILSTHLDABELLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRHLSGYYLTLVKARLPLTTNEKADTMDEGSDVTRQKNGSQSRVCTPQOLLA 1080
Db 1021 SPLFLRHLSGYYLTLVKARLPLTTNEKADTMDEGSDVTRQKNGSQSRVCTPQOLLA 1080
Qy 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRIAELRLTGYSIDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRIAELRLTGYSIDTSLE 1140
Qy 1141 EIFLKVVEECAADTMDEGSCGQHLCTGIAGLDVTLRLKMPPTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTMDEGSCGQHLCTGIAGLDVTLRLKMPPTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARSRRGLFAQIVLPALFVGLALVFSILV 1260
Db 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARSRRGLFAQIVLPALFVGLALVFSILV 1260
Qy 1261 PPFCHYPALRLSPMYGAQVSFFSDEAPGDPGRARLLEALIQEAGLEBPVQHSSHRFSA 1320
Db 1261 PPFCHYPALRLSPMYGAQVSFFSDEAPGDPGRARLLEALIQEAGLEBPVQHSSHRFSA 1320
Qy 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSVWQ 1380
Db 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSVWQ 1380
Qy 1381 NLTCGRNLSDFLVKTYPRLVQGLTKKWVNEVRVGGSLGRDPLSGOELGRSVBELW 1440
Db 1381 NLTCGRNLSDFLVKTYPRLVQGLTKKWVNEVRVGGSLGRDPLSGOELGRSVBELW 1440
Qy 1441 ALLSPLFGALDRVLKNIITAWHSLDAQDSLKIWFNNKGWHSWVAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLFGALDRVLKNIITAWHSLDAQDSLKIWFNNKGWHSWVAFVNRASNAILRAHLP 1500

Qy 1501 PGPARHAHSITTLNHPNLNLTKEQLSEAAIMASSVDVLVSTCVVFAMSFPASFTVLILIE 1560
Db 1501 PGPARHAHSITTLNHPNLNLTKEQLFEAALIMASSVDVLVSTCVVFAMSFPASFTVLILIE 1560
Qy 1561 RVTRAKHLQMLGGLSPTLYWLGNFMDMCMNYLPACIVVLIIFLAFOQRAVAPANLPALL 1620
Db 1561 RVTRAKHLQMLGGLSPTLYWLGNFMDMCMNYLPACIVVLIIFLAFOQRAVAPANLPALL 1620
Qy 1621 LLLLLYGSWTIPMYPASFPFSPSTAYVVLTCINLPIGINGSMATVLELFSQKLOEV 1680
Db 1621 LLLLLYGSWTIPMYPASFPFSPSTAYVVLTCINLPIGINGSMATVLELFSQKLOEV 1680
Qy 1681 SRLKQVFLIIPPHFCILGRGLIDMVRNQAMADAFERLGDROFQSPLRWVVGNLAWVIQ 1740
Db 1681 SRLKQVFLIIPPHFCILGRGLIDMVRNQAMADAFERLGDROFQSPLRWVVGNLAWVIQ 1740
Qy 1741 GPLFLFTLLLOHRSOLLPOPRVRSLLPLGCEEDVVARERERVVQAGTQGDVLVRLNLT 1800
Db 1741 GPLFLFTLLLOHRSOLLPOPRVRSLLPLGCEEDVVARERERVVQAGTQGDVLVRLNLT 1800
Qy 1801 VYRGQMPAVDRICLGIIPGCEFCGLLVNGAGTKTFRMTVGTDLASRGSAVLGHSVAR 1860
Db 1801 VYRGQMPAVDRICLGIIPGCEFCGLLVNGAGTKTFRMTVGTDLASRGSAVLGHSVAR 1860
Qy 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1920
Db 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1920
Qy 1921 RPAGTYSGGNKRKLATALVGPVAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
Db 1921 RPAGTYSGGNKRKLATALVGPVAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
Qy 1981 TSHSMECEALCRLAIMVNGFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAPVAA 2040
Db 1981 TSHSMECEALCRLAIMVNGFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAPVAA 2040
Qy 2041 EFGSELREAHGRLRFLPPLPGRCALARVFGELAVHGAEGHVEDFSVOTMLEEVELYF 2100
Db 2041 EFGSELREAHGRLRFLPPLPGRCALARVFGELAVHGAEGHVEDFSVOTMLEEVELYF 2100
Qy 2101 SKQQKDEDETEQKEAGVGVDPAQGLQHPKRVSQFLDDPSTAEITVL 2146
Db 2101 SKQQKDEDETEQKEAGVGVDPAQGLQHPKRVSQFLDDPSTAEITVL 2146

RESULT 6

AAU09174

ID AAU09174 standard; protein; 2144 AA.

AC AAU09174;

XX 09-APR-2002 (first entry)

DT Human transporter molecule, MTP-1.

XX Human; transporter molecule; MTP-1; cytostatic; Nootropic; HIV;

XX Neuroprotective; Antiparkinsonian; Anticonvulsant; Antianaemic; protein;

KW Antidiabetic; Antiarteriosclerotic; Anti-human immunodeficiency virus;

KW Antiarthritic; Immunosuppressive; Antiashtmatic; Tuberculostatic;

KW Antiulcer; Antimanic; Tranquilliser; Vasotrophic; fertility disorder;

KW transporter-associated disorder; haematopoietic disorder; anxiety;

KW leukocytic disorder; lipid metabolism; abnormal vascularisation;

KW immunological disorder; inflammatory disease; neurological disorder;

KW obsessive-compulsive disorder; cardiac-related disorder;

KW hormonal disorder; reproductive disorder.

XX Homo sapiens.

OS WO200187978-A2.

FN 22-NOV-2001.

XX

XX

XX

PF	14-MAY-2001; 2001WO-US015533.	QY	481	DRFWDGPAADPLTDLRYVMGFFVYLDLVERAAVRVLSGANPRAGLYLQMPYCYVDD	540
XX					
PR	12-MAY-2000; 2000US-0204211P.	DB	479	DRFWDGPAADPLTDLRYVMGFFVYLDLVERAAVRVLSGANPRAGLYLQMPYCYVDD	538
XX					
PA	(MILL-) MILLENNIUM PHARM INC.	QY	541	VFLRVLSRLSPLFLTLAWIYSVTTLTKAVVREKETRLDRDTRAMGLSRAVLWGLFSLCL	600
XX					
PI	Glucksman M, Curtis RAJ;	DB	539	VFLRVLSRLSPLFLTLAWIYSVTTLTKAVVREKETRLDRDTRAMGLSRAVLWGLFSLCL	598
XX					
DR	WPI; 2002-082985/11.	QY	601	GFLLLSAALLVLKIGDILPYSHPGVVFLLAFAVATVTQSPLLSAFFSRANLAAACG	660
XX					
DR	N-PSDB; AAS19207.	DB	599	GFLLLSAALLVLKIGDILPYSHPGVVFLLAFAVATVTQSPLLSAFFSRANLAAACG	658
XX					
PT	New membrane transport protein and polynucleotides, useful for diagnosing	QY	661	GLAYFSLYPLVLCVAMRDRLPAGRVAASLLSPVAFGCGCESLALLBEQGEAGQWHNVG	720
PT	and treating transport protein related disorders e.g. cancer, restenosis,				
PT	asthma and Alzheimer's disease and to identify modulators of therapeutic	DB	659	GLAYFSLYPLVLCVAMRDRLPAGRVAASLLSPVAFGCGCESLALLBEQGEAGQWHNVG	718
XX					
PS	Claim 12; Fig 1; 141pp; English.	QY	721	TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPMNFFRFRSYWCGPRPP	780
XX					
CC	The invention relates to an isolated membrane transporter protein-1 (MTP-	DB	719	TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPMNFFRFRSYWCGPRPP	778
CC	1) (I). (I) is useful for identifying a compound which modulates the				
CC	activity of (I). The method comprises contacting (I) or cell expressing	QY	781	KSPAPCTPLDPKVLVEAPPLGSPGVSRLSKRFPQSPALRGLSLDFVQGHITAPL	840
CC	(I) with a test compound and determining whether (I) binds to the test				
CC	compound or determining the effect of the compound on the activity or	DB	779	KSPAPCTPLDPKVLVEAPPLGSPGVSRLSKRFPQSPALRGLSLDFVQGHITAPL	838
CC	expression of (I). The identified compound is useful in treatment and				
CC	diagnosis of a subject having disorders characterised by aberrant or	QY	841	GHNGAGKTTTLLSGLFPPSGGSFAILGHDRVSRSSMAAIRPHLGVCPOYNVLFDMLTVD	900
CC	unwanted MTP-1 protein or nucleic acid expression or activity, where	DB	839	GHNGAGKTTTLLSGLFPPSGGSFAILGHDRVSRSSMAAIRPHLGVCPOYNVLFDMLTVD	898
CC	transporter-associated disorders include haematopoietic disorders,				
CC	leukocytic disorders, disorders related to lipid metabolism, disorders	QY	901	HWFYGRKLGLSAAVVGPEQRLLQDVLGSKQSVQTRHLSGGMORKLSVAIAFVGGSQV	960
CC	involving abnormal vascularisation, immunological disorders, inflammatory	DB	899	HWFYGRKLGLSAAVVGPEQRLLQDVLGSKQSVQTRHLSGGMORKLSVAIAFVGGSQV	958
CC	diseases, neurological disorders, anxiety disorders, obsessive-compulsive				
CC	disorders, cardiac-related disorders. Disorders also include cellular	QY	961	VILDEPTAGVDPASRRGIWELLKYRGRTLLSTHLDLDEALLGDRVAVVAGRLCCCG	1020
CC	proliferation, growth, differentiation, hormonal disorders and	DB	959	VILDEPTAGVDPASRRGIWELLKYRGRTLLSTHLDLDEALLGDRVAVVAGRLCCCG	1018
CC	reproductive or fertility disorders. The present sequence represents the				
CC	amino acid sequence of human transporter molecule, MTP-1	QY	1021	SPLFRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDTRQEKNGSQSGSRVGTPLLA	1080
XX					
SQ	Sequence 2144 AA;	DB	1019	SPLFRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDTRQEKNGSQSGSRVGTPLLA	1078
QY	Query Match 99.7%; Score 11107; DB 5; Length 2144;	QY	1081	LVQHWVPGARLVEBELPHELVLVLPYTGADHDSFATLRELDTRLAELRLTGYGSDTSL	1140
DB	Best Local Similarity 99.8%; Pred. No. 0;	DB	1079	LVQHWVPGARLVEBELPHELVLVLPYTGADHDSFATLRELDTRLAELRLTGYGSDTSL	1138
	Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;				
QY	1 MAFWTQMLLWKNFMYRRQPVQLLVLLWPLFLFFILVAVRSHHPLEHCHFPNKP 60	QY	1141	EIFLKVVEECAADTMDGSCGQHLCTGIAGLDVTLLKMPPOETALENGSPAGAPETD	1200
DB	1 MAFWTQMLLWKNFMYRRQPVQLLVLLWPLFLFFILVAVRSHHPLEHCHFPNKP 58	DB	1139	EIFLKVVEECAADTMDGSCGQHLCTGIAGLDVTLLKMPPOETALENGSPAGAPETD	1198
QY	61 LPSAGTVPMLOGLICNVNNTCPQPTPGEEPCRLSNFNDLSVRLADARTVLGGASAH 120	QY	1201	QSGSPDAVGRVQGWALTRQQLQALLKRLARSRRLGFAQIVLPALFVGLALVFSLIV	1260
DB	59 LPSAGTVPMLOGLICNVNNTCPQPTPGEEPCRLSNFNDLSVRLADARTVLGGASAH 118	DB	1199	QSGSPDAVGRVQGWALTRQQLQALLKRLARSRRLGFAQIVLPALFVGLALVFSLIV	1258
QY	121 TLAGLGKLIATLRAARSTAAQPOPTKQSPLEPPMLDVAELLTSLRTESLGLALGOAORPL 180	QY	1261	PPFGHYPALRLSPWYGAQVSFFSEDAPGDPGRARLLBALLOEAGLEPPVQHSRHSFA	1320
DB	119 TLAGLGKLIATLRAARSTAAQPOPTKQSPLEPPMLDVAELLTSLRTESLGLALGOAORPL 178	DB	1259	PPFGHYPALRLSPWYGAQVSFFSEDAPGDPGRARLLBALLOEAGLEPPVQHSRHSFA	1318
QY	181 HSLLEAEDLAQELLALSLVELRALLOQRGTSGPLELLSEALCSVRGSPSTVGPSLNW 240	QY	1321	PEVPAEAKVLASGNWTPESPSPACQSCOPCARLLPDCPAAAGPPPPQAVTSGGEVVO	1380
DB	179 HSLLEAEDLAQELLALSLVELRALLOQRGTSGPLELLSEALCSVRGSPSTVGPSLNW 238	DB	1319	PEVPAEAKVLASGNWTPESPSPACQSCOPCARLLPDCPAAAGPPPPQAVTSGGEVVO	1378
QY	241 YEASDLMELVGQEPESALPDGSLSPACSELICALDHPRLSLWRLKPLILGKLLFAPD 300	QY	1381	NLTGRNLSDFLVKTYPRILVRQGLTKKWNVRYGGFSLGGRDPGLPSQBLGRSVEELW	1440
DB	239 YEASDLMELVGQEPESALPDGSLSPACSELICALDHPRLSLWRLKPLILGKLLFAPD 298	DB	1379	NLTGRNLSDFLVKTYPRILVRQGLTKKWNVRYGGFSLGGRDPGLPSQBLGRSVEELW	1438
QY	301 TPFTKRLMAQVNRTEELTLRLDRVEMVEMLGPRIFTFPMNDSSNVAMLQRLQWDEGRR 360	QY	1441	ALLSPLPGGALDRVLKNTAWAHSIDQDSLSKIWFNNKGWHSMAFVNRASNAILRAHL	1500
DB	299 TPFTKRLMAQVNRTEELTLRLDRVEMVEMLGPRIFTFPMNDSSNVAMLQRLQWDEGRR 358	DB	1439	ALLSPLPGGALDRVLKNTAWAHSIDQDSLSKIWFNNKGWHSMAFVNRASNAILRAHL	1498
QY	361 QPRPGRDHMEALRSFLDPGSGGYGQDAHADVGLVGTGLGRVTECLSLDKLEAPSEAA 420	QY	1501	PGPARHAHSITTAHPLNLTKEQLSEALMASSVDVLVSIICVVFAMSFPVASFVLVLEE	1560
DB	359 QPRPGRDHMEALRSFLDPGSGGYGQDAHADVGLVGTGLGRVTECLSLDKLEAPSEAA 418	DB	1499	PGPARHAHSITTAHPLNLTKEQLSEALMASSVDVLVSIICVVFAMSFPVASFVLVLEE	1558
QY	421 LVSRALQLLAEHRFWAGVVFVLPEDSDPTTEHPTDPLGPHVRKIRMDIDVTRTNKIR 480	QY	1561	RVTRAKHLQMLGSLPTLLYMLGNFLWDMCNLYLPACIVVLIFFLAQQRAYVAPANLPALL	1620
DB	419 LVSRALQLLAEHRFWAGVVFVLPEDSDPTTEHPTDPLGPHVRKIRMDIDVTRTNKIR 478				

Db 1559 RVTRAKHLQLMGGLSPTLYLWLNGLFMDMNCVLYVPACIVVLIFFLAQQRAYVAPANLPALL 1618
Qy 1621 LLLLLYKWSITPLMYPASFFPSVSTAYVUTCNINLFIGINGSMTATVLELFSQKLOEV 1680
Db 1619 LLLLLYKWSITPLMYPASFFPSVSTAYVUTCNINLFIGINGSMTATVLELFSQKLOEV 1678
Qy 1681 SRIILKQVFLIPFHCICGLGLDMVRNQMADAFAERLGDQFQSPRLRWEVVGKNNLLAMVIO 1740
Db 1679 SRIILKQVFLIPFHCICGLGLDMVRNQMADAFAERLGDQFQSPRLRWEVVGKNNLLAMVIO 1738
Qy 1741 GPFLLFTLLLOHRSQLLPQPRVRSPLPLGCEDEDAVARERERVVQGGATQGDVLRNLTK 1800
Db 1739 GPFLLFTLLLOHRSQLLPQPRVRSPLPLGCEDEDAVARERERVVQGGATQGDVLRNLTK 1798
Qy 1801 VYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAR 1860
Db 1799 VYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAR 1858
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAGTAGSLARLGSWYAD 1920
Db 1859 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAGTAGSLARLGSWYAD 1918
Qy 1921 RPACTYSGGNKRLATATLALVGDPAVVFLEPTTGMDPSARRFLWNSILLAVVREGRSVML 1980
Db 1919 RPACTYSGGNKRLATATLALVGDPAVVFLEPTTGMDPSARRFLWNSILLAVVREGRSVML 1978
Qy 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Db 1979 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2038
Qy 2041 EFPSELRHAHGRRLRQLPGGRCALARVGEALVHGAHGVEDFSVQTMLEEVFLYF 2100
Db 2039 EFPSELRHAHGRRLRQLPGGRCALARVGEALVHGAHGVEDFSVQTMLEEVFLYF 2098
Qy 2101 SKDGKDEDEEQKEAGVGDVPAPGLQHPKRVSOFLDDPSTAEIVL 2146
Db 2099 SKDGKDEDEEQKEAGVGDVPAPGLQHPKRVSOFLDDPSTAEIVL 2144

RESULT 7

ADD37429
ID ADD37429 standard; protein; 2144 AA.
XX
AC ADD37429;
XX
DT 15-JAN-2004 (first entry)
DE Human transporter MTP-1.
XX
KW Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
KW diabetes.
XX
OS Homo sapiens.
XX
PN US2003143675-A1.
XX
PD 31-JUL-2003.
XX
PF 22-MAY-2002; 2002US-00154419.
XX
PR 12-MAY-2000; 2000US-0204211P.
PR 29-JUN-2000; 2000US-0215376P.
PR 31-JUL-2000; 2000US-0221769P.
PR 19-SEP-2000; 2000US-0233790P.
PR 25-SEP-2000; 2000US-0235107P.
PR 05-OCT-2000; 2000US-0238336P.
PR 14-NOV-2000; 2000US-0248364P.
PR 15-DEC-2000; 2000US-0256240P.
PR 18-DEC-2000; 2000US-0256588P.
PR 21-DEC-2000; 2000US-0258028P.

PR 22-JAN-2001; 2001US-0263169P.
PR 14-MAY-2001; 2001US-00858194.
PR 29-JUN-2001; 2001US-00895811.
PR 31-JUL-2001; 2001US-00919781.
PR 19-SEP-2001; 2001US-00957664.
PR 25-SEP-2001; 2001US-00964295.
PR 05-OCT-2001; 2001US-00972724.
PR 14-NOV-2001; 2001US-00002769.
PR 17-DEC-2001; 2001US-00024623.
PR 22-JAN-2002; 2002US-00055025.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Glucksmann MA, Meyers RE;
XX
DR WFI; 2003-851783/79.
DR N-PSDB; ADD37428, ADD37430.
XX
PT New isolated nucleic acid, useful for preparing a composition for
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
PT or diabetes.
XX
PS Claim 11; SEQ ID NO 2; 663pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.
XX
SQ Sequence 2144 AA;

Query Match 99.7%; Score 11107; DB 7; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 MAFWTQLMLLLKWNFMYRRRQPVQLLVLELWPLFFLFFILVAVRHSHPPLHEHCHFFNKP 60
Db 1 MAFWTQLMLLLKWNFMYRRRQPVQLLVLELWPLFFLFFILVAVRHSHPPLHEHCHFFNKP 58
Qy 61 LPSAGTVPMLOGLICNVNNTCFPQTPGHEPGRHSNFDNSLVSRLLADARTVLGGASAH 120
Db 59 LPSAGTVPMLOGLICNVNNTCFPQTPGHEPGRHSNFDNSLVSRLLADARTVLGGASAH 118
Qy 121 TLAGLGLKLTATLRAARSTAQPTKOSPPLPPMMDVAELLTSLRTESLGLAQOAEPL 180
Db 119 TLAGLGLKLTATLRAARSTAQPTKOSPPLPPMMDVAELLTSLRTESLGLAQOAEPL 178
Qy 181 HSLLAEADLAQELLALRSILVELRALLQRPRTSGPGLSEALCSVRGSPSTVGPSSLNW 240
Db 179 HSLLAEADLAQELLALRSILVELRALLQRPRTSGPGLSEALCSVRGSPSTVGPSSLNW 238
Qy 241 YEASDLMELVQEPESALPDSSLSACSELIIGALDSHPISRLLRRLKPLILKLLFAPD 300
Db 239 YEASDLMELVQEPESALPDSSLSACSELIIGALDSHPISRLLRRLKPLILKLLFAPD 298
Qy 301 TPETKLMQVNTFEELTLRDVREVMELGPRIETFMNDSSNVAMQLQLQMDGRR 360
Db 299 TPETKLMQVNTFEELTLRDVREVMELGPRIETFMNDSSNVAMQLQLQMDGRR 358
Qy 361 QRPFGGRDHMEALRSFLDPGSGGYGQDAHADVGHVLTIGRVTECLSLDKLEAAPSEAA 420
Db 359 QRPFGGRDHMEALRSFLDPGSGGYGQDAHADVGHVLTIGRVTECLSLDKLEAAPSEAA 418

QY 421 LVSRALQLAEHRFWAGVFLGPEDSSDPTEHPTDLGFGHVRIRKIMDIDVVTNKR 480
DB 419 LVSRALQLAEHRFWAGVFLGPEDSSDPTEHPTDLGFGHVRIRKIMDIDVVTNKR 478
QY 481 DRFPDGPAAADPLTDLRYVWGFFVYLODLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
DB 479 DRFPDGPAAADPLTDLRYVWGFFVYLODLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 538
QY 541 VFLRVLRSPLPLFLTLAWIYSVTLVKAVVREKETRLRDTMRAMGLSRAVLWGLFSLCL 600
DB 539 VFLRVLRSPLPLFLTLAWIYSVTLVKAVVREKETRLRDTMRAMGLSRAVLWGLFSLCL 598
QY 601 GPFLLSAALLVLVLKGLDILPSHPGVVFLFLAFAFVATVTSQFLLSAFFSRANLAAACG 660
DB 599 GPFLLSAALLVLVLKGLDILPSHPGVVFLFLAFAFVATVTSQFLLSAFFSRANLAAACG 658
QY 661 GLAYFSLYLPYVLCVAMDRDLPAGGRVAASLLSPVAFGCBESLALLEEQEGEAQHNVG 720
DB 659 GLAYFSLYLPYVLCVAMDRDLPAGGRVAASLLSPVAFGCBESLALLEEQEGEAQHNVG 718
QY 721 TRPTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPWNFPFRSYWCGPRPP 780
DB 719 TRPTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPEPWNFPFRSYWCGPRPP 778
QY 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPALRGLSLDFYQGHITAF 840
DB 779 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPALRGLSLDFYQGHITAF 838
QY 841 GHNGAGKTTTILSGLPFPSSGSAFIIIGHDVSRSSMAAIRPHLGVCPQYNVLFMDLTVD 900
DB 839 GHNGAGKTTTILSGLPFPSSGSAFIIIGHDVSRSSMAAIRPHLGVCPQYNVLFMDLTVD 898
QY 901 HWVFYGRUKLSAAVVGPEORLLQDVLVSKQSVQTRHLSGGMORKLSVAIAFVGGSOV 960
DB 899 HWVFYGRUKLSAAVVGPEORLLQDVLVSKQSVQTRHLSGGMORKLSVAIAFVGGSOV 958
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1020
DB 959 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1018
QY 1021 SPLFRRHLGSGYILTUKARLPLTTNEKADTMEGSDVTRQEKNGQSGSRVGTPLL 1080
DB 1019 SPLFRRHLGSGYILTUKARLPLTTNEKADTMEGSDVTRQEKNGQSGSRVGTPLL 1078
QY 1081 LVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLPELDTLAEALRTGVTSDTSL 1140
DB 1079 LVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLPELDTLAEALRTGVTSDTSL 1138
QY 1141 EIFLKVVECAADTDMEDSGCGHLCTGIAGLDVTLRLKMPPEQETALENGEPAGSAPETD 1200
DB 1139 EIFLKVVECAADTDMEDSGCGHLCTGIAGLDVTLRLKMPPEQETALENGEPAGSAPETD 1198
QY 1201 QGSGPDVAVRQGWALTQOQLALLKFLPLARRSRGLPFAQIVLPALFVGLALVFSILV 1260
DB 1199 QGSGPDVAVRQGWALTQOQLALLKFLPLARRSRGLPFAQIVLPALFVGLALVFSILV 1258
QY 1261 PFGHYPALRLSPPTYGAGVFFESDAPDPCRAFLLEALLQEAGLEPPPVQHSHPESA 1320
DB 1259 PFGHYPALRLSPPTYGAGVFFESDAPDPCRAFLLEALLQEAGLEPPPVQHSHPESA 1318
QY 1321 PEVPAEAVKVLASGNWTPESPSPACQSPQGARILLPCPAAAGPPPPQAVTSGSEVVQ 1380
DB 1319 PEVPAEAVKVLASGNWTPESPSPACQSPQGARILLPCPAAAGPPPPQAVTSGSEVVQ 1378
QY 1381 NLTRNLSDFLVKTYPRLVRQGLKTKWNVVYVGGFSLGGDRDPLGSGELGRSVEELW 1440
DB 1379 NLTRNLSDFLVKTYPRLVRQGLKTKWNVVYVGGFSLGGDRDPLGSGELGRSVEELW 1438
QY 1441 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSMAFVFNRSNAIILRAHLP 1500
DB 1439 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSMAFVFNRSNAIILRAHLP 1498

QY 1501 PGPAPHAHSITTLNHNPLNLTKEQLSEAAALMASSVDVLVSYCVVFAMSFVPASFTLVLIB 1560
DB 1499 PGPAPHAHSITTLNHNPLNLTKEQLSEAAALMASSVDVLVSYCVVFAMSFVPASFTLVLIB 1558
QY 1561 RVTRAKHLQLMGGSLPTLYWLGNFWDNMCNLYLPACIVVLIPAFQORAYVAPANLPALL 1620
DB 1559 RVTRAKHLQLMGGSLPTLYWLGNFWDNMCNLYLPACIVVLIPAFQORAYVAPANLPALL 1618
QY 1621 LLLLYLWHSITPLMYPASFFSVSTAVVLTICNLFIINGSMATFVLELPSDOKLQSV 1680
DB 1619 LLLLYLWHSITPLMYPASFFSVSTAVVLTICNLFIINGSMATFVLELPSDOKLQSV 1678
QY 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPURWEVGNKLLAMVTQ 1740
DB 1679 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPURWEVGNKLLAMVTQ 1738
QY 1741 GFLFLLFTLLLOHRSQLLPQPRVRSPLPLGDEEDVARERERVQATQGDVILVRLNLT 1800
DB 1739 GFLFLLFTLLLOHRSQLLPQPRVRSPLPLGDEEDVARERERVQATQGDVILVRLNLT 1798
QY 1801 VYRGQMPAVDRLCGLIPPGECFGLGVNGAGKTSTFRMVTGDTLASRGEAVLAGHSVAR 1860
DB 1799 VYRGQMPAVDRLCGLIPPGECFGLGVNGAGKTSTFRMVTGDTLASRGEAVLAGHSVAR 1858
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVATAGSGLARLGLSWYAD 1920
DB 1859 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVATAGSGLARLGLSWYAD 1918
QY 1921 RPAGTYSGGNKRLATALLVGDPAVFLDPTTGMDSARRFLMNSLLAVVREGRSVNL 1980
DB 1919 RPAGTYSGGNKRLATALLVGDPAVFLDPTTGMDSARRFLMNSLLAVVREGRSVNL 1978
QY 1981 TSHSMECEALCSRLATMNGRFRCLGSPQHLKGRFAAGHTLTLLRVPAAARQOPAAAFVAA 2040
DB 1979 TSHSMECEALCSRLATMNGRFRCLGSPQHLKGRFAAGHTLTLLRVPAAARQOPAAAFVAA 2038
QY 2041 EFPSELREAHGRLRQLPFGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
DB 2039 EFPSELREAHGRLRQLPFGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2098
QY 2101 SKDQKQEDTBEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146
DB 2099 SKDQKQEDTBEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2144

RESULT 8
ABUS4629
ID ABUS4629 standard; protein; 2059 AA.
XX
AC ABUS4629;
XX
DT 03-JUN-2003 (first entry)
XX
DE Human NOVX polypeptide #88.
XX
KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerositis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS Homo sapiens.
XX
FN WO200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.

PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 06-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294184P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-03322115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigar M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
XX WPI: 2003-046858/04.
DR N-PSDB; ABX72257.
XX
PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 281-282; 666pp; English.
XX
CC The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atriocentricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberos sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC hematopoietic disorders, haemophilia, hypercoagulation. Crohn's disease
CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
CC of the invention
XX
SQ Sequence 2059 AA;
XX

Query Match 95.5%; Score 10639.5; DB 6; Length 2059;

Best Local Similarity 95.9%; Pred. No. 0;
Matches 2058; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 MAFWTQMLLLWKNFMYRRRQPVQLLVLLWPLFLFFILVAVRSHHPPLSHHCHFFNKP 60
Db 1 MAFWTQMLLLWKNFMYRRRQPVQLLVLLWPLFLFFILVAVRSHHPPLSHHCHFFNKP 60
Qy 61 LPSAGTVPWLOGLICNVNNTCFPOLTPGEBPGRLSNFNDLSVRLLDADATVILGASAH 120
Db 61 LPSAGTVPWLOGLICNVNNTCFPOLTPGEBPGRLSNFNDLSVRLLDADATVILGASAH 120
Qy 121 TLAGLGLKIATLAARSTAOPTKOSPPLPMLDVAELTSLRTESLGLALQAOEPL 180
Db 121 TLAGLGLKIATLAARSTAOPTKOSPPLPMLDVAELTSLRTESLGLALQAOEPL 180
Qy 181 HSLLEAAEDLAQELLALRSIVELRALLQRPRTSGPLELSEALCSVRGSSSTVGP SLNW 240
Db 181 HSLLEAAEDLAQELLALRSIVELRALLQRPRTSGPLELSEALCSVRGSSSTVGP SLNW 240
Qy 241 YEASDLMELVQBPESALPDSSLPACSELI GALDSHPILSRLLWRRLKPIILGKLLFAPD 300
Db 241 YEASDLMELVQBPESALPDSSLPACSELI GALDSHPILSRLLWRRLKPIILGKLLFAPD 300
Qy 301 TPTRKLMQVNRTPFELTLLRDVREVMGLGPRIFTFMNDSNVAMLQRLLOMQDEGR 360
Db 301 TPTRKLMQVNRTPFELTLLRDVREVMGLGPRIFTFMNDSNVAMLQRLLOMQDEGR 360
Qy 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
Db 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 420
Qy 421 LVSRAQLLAHHRFWAGVFLGPEDSDPTHEHTPDLPDGPCHVRIKIRMDIVVTRTNKIR 480
Db 421 LVSRAQLLAHHRFWAGVFLGPEDSDPTHEHTPDLPDGPCHVRIKIRMDIVVTRTNKIR 480
Qy 481 DRFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAGILYLOMPYPCVDD 540
Db 481 DRFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAGILYLOMPYPCVDD 540
Qy 541 VFLRVLRSRLPLFLTLLAWIYSVTLTVKA VVREKTRLRDTMRAMGLSRAVLWGLFSLCL 600
Db 541 VFLRVLRSRLPLFLTLLAWIYSVTLTVKA VVREKTRLRDTMRAMGLSRAVLWGLFSLCL 600
Qy 601 GPFLLSNALLVILKGDILIPYSHPGVFLFAAFAVATVTSQFLLSAFAFSRANLAACG 660
Db 601 GPFLLSNALLVILKGDILIPYSHPGVFLFAAFAVATVTSQFLLSAFAFSRANLAACG 660
Qy 661 GLAYFSLYLPVLCVAVWRDLRDPAGGRVAASLLSPVAFPGCESIALILEQEGEQAHNVG 720
Db 661 GLAYFSLYLPVLCVAVWRDLRDPAGGRVAASLLSPVAFPGCESIALILEQEGEQAHNVG 720
Qy 721 TRPTADVLSLAQVSGLLLDALYGLATWYLEAVCPQGYGIPFPWNPFFRRSYWCGRRPP 780
Db 721 TRPTADVLSLAQVSGLLLDALYGLATWYLEAVCPQGYGIPFPWNPFFRRSYWCGRRPP 780
Qy 781 KSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRPSPQALRGILSLDFYQGHITAF 840
Db 781 KSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRPSPQALRGILSLDFYQGHITAF 840
Qy 841 GHNGAGKTTTSLISGLFPSPGSAFTILGHDVRSMAAIPHLGVCQYQNVLPDMLTVDE 900
Db 841 GHNGAGKTTTSLISGLFPSPGSAFTILGHDVRSMAAIPHLGVCQYQNVLPDMLTVDE 900
Qy 901 HVMFYGLKGLSAAVVGPEODRLLDQVGLSVKSVQTRHLSGGMQRLKSLVAIAFVGSQV 960
Db 901 HVMFYGLKGLSAAVVGPEODRLLDQVGLSVKSVQTRHLSGGMQRLKSLVAIAFVGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLAKYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLAKYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGYYLTILVKARLPLTNEKADTDMEGSVDTROBKNGSQSRSVGTPOLLA 1080
Db 1021 SPLFLRRHLGSGYYLTILVKARLPLTNEKADTDMEGSVDTROBKNGSQSRSVGTPOLLA 1080

Db 1021 SPFLFRRHLSGYYTLVKARLPLTTNEKADTMEGSDVTRQEKXGSGSRVGTFFQLA 1080
Qy 1081 LVQHWVPGARLVEELPHELVLVLVLYTGAHDGSPATLFRLEDTRLAELRLTGYGIDTSL 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLVLYTGAHDGSPATLFRLEDTRLAELRLTGYGIDTSL 1140
Qy 1141 EFLFKVBECAADTMEDGSCQHLCCTGTAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
Db 1141 EFLFKVBECAADTMEDGSCQHLCCTGTAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
Qy 1201 QSGGPDVAVRGVGMWLTROOLQALLKRELLARRRRGLFAQIVLPALFVGLALVPSLIV 1260
Db 1201 QSGGPDVAVRGVGMWLTROOLQALLKRELLARRRRGLFAQIVLPALFVGLALVPSLIV 1260
Qy 1261 PPFHYPALRLSPTMYGAQVSFFSADAPGDPGRARLLEALLQEALEPPVQSHSRFSA 1320
Db 1261 PPFHYPALRLSPTMYGAQVSFFSADAPGDPGRARLLEALLQEALEPPVQSHSRFSA 1320
Qy 1321 PEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGEVVQ 1380
Db 1321 PEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGEVVQ 1380
Qy 1381 NLTGRNLSDPLVKTYPRLVRQGLTKKWNVEYVGGFSLGGDRDGLPSGOELGRSVEBLW 1440
Db 1381 NLTGRNLSDPLVKTYPRLVRQGLTKKWNVEYVGGFSLGGDRDGLPSGOELGRSVEBLW 1440
Qy 1441 ALLSPLPGGALDRVLKNTAWAHSDDAODSLKIFNNKGWHSWAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSDDAODSLKIFNNKGWHSWAFVNRASNAILRAHLP 1500
Qy 1501 PGPARHAHSITTLNHLPLNLTKEQLSEAAALMASSVDVLSICVVFAMSFVPASFTLVLEE 1560
Db 1501 PGPARHAHSITTLNHLPLNLTKEQLSEAAALMASSVDVLSICVVFAMSFVPASFTLVLEE 1560
Qy 1561 RVTRAKHLQLMGGLSFTLYWLNFLWDMCNVYLPACIVVLIFLAFQQRAYVAPANLPALL 1620
Db 1561 RVTRAKHLQLMGGLSFTLYWLNFLWDM----- 1588
Qy 1621 LLLLLYGSITPLMYPASFPFSPVSTAYVLTCTINLFIGINGSMATFVLELPSDKLOEV 1680
Db 1589 -----KLQEV 1593
Qy 1681 SRILKQVFLIPHPCLGRGLDMVRNQAMADAFERLGRDQFQSPLRWEVVGKLLAMVIQ 1740
Db 1594 SRILKQVFLIPHPCLGRGLDMVRNQAMADAFERLGRDQFQSPLRWEVVGKLLAMVIQ 1653
Qy 1741 GPLFLLFTLLQHRSQLLPQPRVRSLLPILGEDEDVARERERVQATQGDVVLRLNLT 1800
Db 1654 GPLFLLFTLLQHRSQLLPQPRVRSLLPILGEDEDVARERERVQATQGDVVLRLNLT 1713
Qy 1801 VYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSTFRWVTGDTLASRGEAVLAGHSVAR 1860
Db 1714 VYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSTFRWVTGDTLASRGEAVLAGHSVAR 1773
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1920
Db 1774 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1833
Qy 1921 RPAGTYSGNKKRLATALLVGDPAVFLDEPTTGMDFPSARRFLWNSLLAVVREGRSVNL 1980
Db 1834 RPAGTYSGNKKRLATALLVGDPAVFLDEPTTGMDFPSARRFLWNSLLAVVREGRSVNL 1893
Qy 1981 TSHSMECEALCSRLAIWNGFRCLGSPQHLKGRFAAGHTITLVRPAARSQPAAFVAA 2040
Db 1894 TSHSMECEALCSRLAIWNGFRCLGSPQHLKGRFAAGHTITLVRPAARSQPAAFVAA 1953
Qy 2041 EFPQSELREAHGRLRFLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
Db 1954 EFPQSELREAHGRLRFLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2013
Qy 2101 SKDQKQEDTBEQKAGVGDPAVGLQHPKRVSRQFLDDPSTAEVTL 2146
Db 2014 SKDQKQEDTBEQKAGVGDPAVGLQHPKRVSRQFLDDPSTAEVTL 2059

RESULT 9
ABU08464
ID ABU08464 standard; protein; 2008 AA.
XX
AC ABU08464;
XX
DT 18-JUN-2003 (first entry)
XX
DE Amino acid sequence for human ABCA7 splice variant #1.
XX
KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
immunomodulator; immunosuppressive; antiinflammatory;
antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 1..28
FT /note="Given as SEQ ID No:9 and specifically claimed in
FT Claim 2"
XX
PN WO2003010315-A1.
XX
XX 06-FEB-2003.
XX
XX 24-JUL-2002; 2002WO-JP007487.
XX
XX 25-JUL-2001; 2001JP-00224176.
XX
XX 06-DEC-2001; 2001JP-00372530.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
XX Ueda K, Nakagawa S, Nagase T;
XX
XX WPI; 2003-239444/23.
XX
XX N-PSDB; ABX95283.
XX
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
the immune system, applicable in diagnosis of and screening drugs for
e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
XX Claim 1; Page 106-115; 183pp; Japanese.
XX
XX The present invention relates to the isolation of human ATP-binding
cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
polynucleotide sequences encoding them. The protein is applicable in the
diagnosis and screening of drugs for autoimmune diseases, Sjogren's
syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
It may also be used in a method for screening ABCA-SSN inhibitors. The
present sequence represents human ABCA7 splice variant #1
XX
SQ Sequence 2008 AA;

Query Match 92.1%; Score 10264; DB 6; Length 2008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 167 ESLGLAGQAQCEPLHSLLEAAEDLAQELLALRSVLVELRALQRPRTSGPLELSEALCS 226
Db 29 ESLGLAGQAQCEPLHSLLEAAEDLAQELLALRSVLVELRALQRPRTSGPLELSEALCS 88
Qy 227 VRGSPSTVGPSSINWYEASDLMLVQBPESALPDSSLSPPACSELTGALDHSPLSLWRR 286
Db 89 VRGSPSTVGPSSINWYEASDLMLVQBPESALPDSSLSPPACSELTGALDHSPLSLWRR 148
Qy 287 LKPLILGKLLFAPDTPFTRKLMQVNRTFEELTLRLDRVREVMELGPRIFTFMNDSSNA 346
Db 149 LKPLILGKLLFAPDTPFTRKLMQVNRTFEELTLRLDRVREVMELGPRIFTFMNDSSNA 208

Qy	347	MLQRLQMDGRRQPPGGEDHMEALRSFLDPGSGGYSWQDAHADYCHLVGTLGRVTEC	406
Db	209	MLQRLQMDGRRQPPGGEDHMEALRSFLDPGSGGYSWQDAHADYCHLVGTLGRVTEC	269
Qy	407	LSLDKLEAAPSEAAALVSRAQLLAHHRFWAGVFLGPEDSSDPTEHPTPDLPQGHVRIKI	466
Db	269	LSLDKLEAAPSEAAALVSRAQLLAHHRFWAGVFLGPEDSSDPTEHPTPDLPQGHVRIKI	328
Qy	467	RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	526
Db	329	RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	388
Qy	527	LYLQOMPYPVCVDDVFLRVLSRSPLFLTLAWIYSVTLTKAVVREKETRLDTRMAMGL	586
Db	389	LYLQOMPYPVCVDDVFLRVLSRSPLFLTLAWIYSVTLTKAVVREKETRLDTRMAMGL	448
Qy	587	SRAVLWLGWFLSCIGPFLLSAALLVLVLKLGDIILPYSHPGVVFLLAAFAVATVTSQFLL	646
Db	449	SRAVLWLGWFLSCIGPFLLSAALLVLVLKLGDIILPYSHPGVVFLLAAFAVATVTSQFLL	508
Qy	647	SAFESRANLAAACGGLAYFSLPYLVCVWRDLRDPAGGRVAASLLSPVAFGFCESLAL	706
Db	509	SAFESRANLAAACGGLAYFSLPYLVCVWRDLRDPAGGRVAASLLSPVAFGFCESLAL	568
Qy	707	LEEQEGEAQHNVTGRTADVFSLAQVSGILLDAAALYGLATWYLEAVCPGQYGIPEPWN	766
Db	569	LEEQEGEAQHNVTGRTADVFSLAQVSGILLDAAALYGLATWYLEAVCPGQYGIPEPWN	628
Qy	767	FPFRRSYWCGRPRPKSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG	826
Db	629	FPFRRSYWCGRPRPKSPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG	688
Qy	827	LSLDFYQCHITAFIAGHNGAGKTTTSLISGLFPPSGGSFALGHVDRSSMAAIRPHLVC	886
Db	689	LSLDFYQCHITAFIAGHNGAGKTTTSLISGLFPPSGGSFALGHVDRSSMAAIRPHLVC	748
Qy	887	PQYNVLFDMLTVDHVMFYGELKGLSAVNGPEODRLLQDGLVSKOSVOTRHLSGMQR	946
Db	749	PQYNVLFDMLTVDHVMFYGELKGLSAVNGPEODRLLQDGLVSKOSVOTRHLSGMQR	808
Qy	947	KLVAIAFVGGSVQVILDEPTAGVDPPASRRGIWELLKLYREGRTLILSTHHLDEALLGD	1006
Db	809	KLVAIAFVGGSVQVILDEPTAGVDPPASRRGIWELLKLYREGRTLILSTHHLDEALLGD	868
Qy	1007	RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTNEKADTDMEGSVDTROBKN	1066
Db	869	RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTNEKADTDMEGSVDTROBKN	928
Qy	1067	GSQSRVGTGQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFRDLDTLAE	1126
Db	929	GSQSRVGTGQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFRDLDTLAE	988
Qy	1127	LRLTGYGISTSLBEEIFLKVVEECAADTDMEDGSCQHLCGTAGLDVTLRLKMPPOETA	1186
Db	989	LRLTGYGISTSLBEEIFLKVVEECAADTDMEDGSCQHLCGTAGLDVTLRLKMPPOETA	1048
Qy	1187	LENCEPAGSAPETDQSGSPDAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFLPAQIVLP	1246
Db	1049	LENCEPAGSAPETDQSGSPDAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFLPAQIVLP	1108
Qy	1247	ALFVGLALVSLIIVPPFGHPALRLSTMTYGAQVSFESEADPGDPGARLLEALLQBAGL	1306
Db	1109	ALFVGLALVSLIIVPPFGHPALRLSTMTYGAQVSFESEADPGDPGARLLEALLQBAGL	1168
Qy	1307	EEPVPQSHSHRFSAPPEVAEVAKVLASGNMTPEPSPACOCOPGARRLLPDCPAAAGGP	1366
Db	1169	EEPVPQSHSHRFSAPPEVAEVAKVLASGNMTPEPSPACOCOPGARRLLPDCPAAAGGP	1228
Qy	1367	PPQAVTGSGEVWQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGGFSLGGRDPGL	1426
Db	1229	PPQAVTGSGEVWQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGGFSLGGRDPGL	1288

Qy	1427	PSQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSIDAQDSLKIWFNNKWHSMVAF	1486
Db	1289	PSQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSIDAQDSLKIWFNNKWHSMVAF	1348
Qy	1487	VNRASNAILRAHLPPOPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM	1546
Db	1349	VNRASNAILRAHLPPOPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM	1408
Qy	1547	SFVPASFTLLIIBERVTRAKHLQIMGLSPTLWLGFLWDMCNLYLPACIVVLIFLAFO	1606
Db	1409	SFVPASFTLLIIBERVTRAKHLQIMGLSPTLWLGFLWDMCNLYLPACIVVLIFLAFO	1468
Qy	1607	QRAVAPANIPALLLLLLYGVWSITPLMYPASFEFSPSTAYVVLTCINLFIGINGSMAT	1666
Db	1469	QRAVAPANIPALLLLLLYGVWSITPLMYPASFEFSPSTAYVVLTCINLFIGINGSMAT	1528
Qy	1667	FVLELFSQDLQLEVSRLIKQVFLIFPHFCLGRGLIDMVRNQANMADAFERLGRDQFQSPLR	1726
Db	1529	FVLELFSQDLQLEVSRLIKQVFLIFPHFCLGRGLIDMVRNQANMADAFERLGRDQFQSPLR	1588
Qy	1727	WEVVGKXLLAMVIOGPELFTLLLOHRSQLLLPQVRSLPILGEEDEDVARERERVQO	1786
Db	1589	WEVVGKXLLAMVIOGPELFTLLLOHRSQLLLPQVRSLPILGEEDEDVARERERVQO	1648
Qy	1787	ATOGDVLVLRNLTGVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFRVMTGDTLA	1846
Db	1649	ATOGDVLVLRNLTGVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFRVMTGDTLA	1708
Qy	1847	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA	1906
Db	1709	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA	1768
Qy	1907	GSGLARLGLSWYADRPAGTVSGGNKRLATALALVGDPAVFLDEPTTGMDDPSARRPLWN	1966
Db	1769	GSGLARLGLSWYADRPAGTVSGGNKRLATALALVGDPAVFLDEPTTGMDDPSARRPLWN	1828
Qy	1967	SLIAVVRGREGSVMLTSHSMECEALCSRLAIWNGFRCLGSPQHLKGRFAAGHTLTLRV	2026
Db	1829	SLIAVVRGREGSVMLTSHSMECEALCSRLAIWNGFRCLGSPQHLKGRFAAGHTLTLRV	1888
Qy	2027	PAARSQPAAFVAAEFPGSELREAHGRLRFQLPQGRCALARVFGELAVHGAEHGVEDF	2086
Db	1889	PAARSQPAAFVAAEFPGSELREAHGRLRFQLPQGRCALARVFGELAVHGAEHGVEDF	1948
Qy	2087	SVSQTMLEEVFLYFSKQDKDETEQKEAGVGDPAQGLQHPKRVSOQLDDPSTAEATVL	2146
Db	1949	SVSQTMLEEVFLYFSKQDKDETEQKEAGVGDPAQGLQHPKRVSOQLDDPSTAEATVL	2008

RESULT 10
ABU08465
ID ABU08465 standard; protein; 1993 AA.
XX AC ABU08465;
XX AC
XX AC
DT 18-JUN-2003 (first entry)
XX
XX Amino acid sequence for human ABCA7 splice variant #2.
DE DE
XX Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
antiarteriosclerotic.
XX
OS Homo sapiens.
XX
XX WO2003010315-A1.
PN
XX
PD 06-FEB-2003.
XX
PF 24-JUL-2002; 2002WO-JP007487.
XX

PR 25-JUL-2001; 2001JP-002241176.
PR 06-DEC-2001; 2001JP-00372530.
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
DR N-PSDB; ABX95284.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 1; Page 145-154; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #2
XX
SQ Sequence 1993 AA;
Query Match 91.3%; Score 10171.5; DB 6; Length 1993;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
QY 167 ESIGLALGQAQEPHLSLEAEADLAQELIALSLVELRALLOPRGTSGPLELLSEALCS 226
DB 29 ESIGLALGQAQEPHLSLEAEADLAQELIALSLVELRALLOPRGTSGPLELLSEALCS 88
QY 227 VRGSPSTVGPSLNWYEASDLMELVQEPESALPDSLSLSPACSELI GALSHPLSLLWRR 286
DB 89 VRGSPSTVGPSLNWYEASDLMELVQEPESALPDSLSLSPACSELI GALSHPLSLLWRR 148
QY 287 LKPLILGKLFPADPTFTTKMAQVNRTPFEELTLARDVREVMELGPRIFTPMNDSSNVA 346
DB 149 LKPLILGKLFPADPTFTTKMAQVNRTPFEELTLARDVREVMELGPRIFTPMNDSSNVA 208
QY 347 MLQRLQMQDEGRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 406
DB 209 MLQRLQMQDEGRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 268
QY 407 LSLDKLEAAPSEALVSRALQLLASHRFWAGVVFILGPDSDDPTSHPTDPLGPHVRIKI 466
DB 269 LSLDKLEAAPSEALVSRALQLLASHRFWAGVVFILGPDSDDPTSHPTDPLGPHVRIKI 328
QY 467 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 526
DB 329 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 388
QY 527 LYLOQMPYPCYVDDVFLVLSRLPLFTLWIVSYVTLTVKAVVREKETRLRDTNRAMGL 586
DB 369 LYLOQMPYPCYVDDVFLVLSRLPLFTLWIVSYVTLTVKAVVREKETRLRDTNRAMGL 448
QY 587 SRAVLWLGWFLSCLGFFLLSALLVLVLKGLDILPSHPGVVFLFAAPAVATVTSFLL 646
DB 449 SRAVLWLGWFLSCLGFFLLSALLVLVLKGLDILPSHPGVVFLFAAPAVATVTSFLL 508
QY 647 SAFFSRANLAACGGLAYFSLYLPVLCVWNRDLPAAGRVAASLLSPVAFGCGESLAL 706
DB 509 SAFFSRANLAACGGLAYFSLYLPVLCVWNRDLPAAGRVAASLLSPVAFGCGESLAL 568
QY 707 LEEQEGEQAHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPQGYGPEPWN 766
DB 569 LEEQEGEQAHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPQGYGPEPWN 628
QY 767 PFFRYSYWGCPRPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRPFGSPOPALRG 826
DB

DB 629 PFFRYSYWGCPRPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRPFGSPOPALRG 688
QY 827 LSLDFYQGHITAFILGHNGAGKTTTSLIISGLPSPGGSFAFIIGHDVRSMAAIRPHLGYC 886
DB 689 LSLDFYQGHITAFILGHNGAGKTTTSLIISGLPSPGGSFAFIIGHDVRSMAAIRPHLGYC 748
QY 887 PQYNVLFDMILTVDHFWFYGRILKGLSAAVVGPEOQRLQDVLGVSKQSVQTRHLSGGMOR 946
DB 749 PQYNVLFDMILTVDHFWFYGRILKGLSAAVVGPEOQRLQDVLGVSKQSVQTRHLSGGMOR 808
QY 947 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREORTILSTHLLDEALLGD 1006
DB 809 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREORTILSTHLLDEALLGD 868
QY 1007 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTIVKARLPLTTNEKADTDMEGSVDTQEKKN 1066
DB 869 RVAVVAGRLCCCGSPLFLRRHLGSGYYLTIVKARLPLTTNEKADTDMEGSVDTQEKKN 928
QY 1067 GSQGSRVGTQQLLALVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLRFELDTRLAE 1126
DB 929 GSQGSRVGTQQLLALVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLRFELDTRLAE 988
QY 1127 LRLTYGISTDSLEBIFLKVVEECAAADTDMEDGSCQHLCCTGIAGLDVTLRLKMPQETA 1186
DB 989 LRLTYGISTDSLEBIFLKVVEECAAADTDMEDGSCQHLCCTGIAGLDVTLRLKMPQETA 1048
QY 1187 LENGEPAGSAPETDQSGPDVAVGRVQGWALTQQLQALLKFFLLARRSRRLGFAQIVLP 1246
DB 1049 LENGEPAGSAPETDQSGPDVAVGRVQGWALTQQLQALLKFFLLARRSRRLGFAQIVLP 1108
QY 1247 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSFFSEDPGDPGRARLLBALQEAQL 1306
DB 1109 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSFFSEDPGDPGRARLLBALQEAQL 1168
QY 1307 BEPPVQHSRHPFAPEVPAEVAKVASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1366
DB 1169 BEPPVQHSRHPFAPEVPAEVAKVASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1228
QY 1367 PPQAVTSGGEVQNLGTGRNLSDFLVKTYPRLVROGLTKKWNVNRYVGGFSLGGRDPL 1426
DB 1229 PPQAVTSGGEVQNLGTGRNLSDFLVKTYPRLVROGLTKKWNVNRYVGGFSLGGRDPL 1288
QY 1427 PSGELGRSVEBELWALLSPLPGGALDRVLKXLTAWAHS LDAQDSLKIFWNNKGHSMVAF 1486
DB 1289 PSGELGRSVEBELWALLSPLPGGALDRVLKXLTAWAHS LDAQDSLKIFWNNKGHSMVAF 1348
QY 1487 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI CVVFAM 1546
DB 1349 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI CVVFAM 1408
QY 1547 SFVPASFVLVIERVTRAKHLQLMGGISPTLYMGNFLWDMCNVLPACI VVLIPLAFQ 1606
DB 1409 SFVPASFVLVIERVTRAKHLQLMGGISPTLYMGNFLWDMCNVLPACI VVLIPLAFQ 1468
QY 1607 QRAYVAPANLALLLLLYGMSITPLMYPASFFFSVPSTAYVLTICNLFIGINGSMAT 1666
DB 1469 QRAYVAPANLALLLLLYGMSITPLMYPASFFFSVPSTAYVLTICNLFIGINGSMAT 1528
QY 1667 FVLELFSQKLQSVRIKQVFLIPPHFCLOGLIDMVNRQAMADAFERLGRDQFQSPUR 1726
DB 1529 FVLELFSQKLQSVRIKQVFLIPPHFCLOGLIDMVNRQAMADAFERLGRDQFQSPUR 1588
QY 1727 WEVVGKNLAMVIOGPFLLFTLLQHSQLLPOPRVRSPLPLGGEDEDDVARERERVQ 1786
DB 1589 WEVVGKNLAMVIOGPFLLFTLLQHSQLLPOPRVRSPLPLGGEDEDDVARERERVQ 1648
QY 1787 ATQGDVLVLRNLTKVYRGORMPAVDRCLGIPPGECFGLLVNGAGKSTSTFRMVTGDTLA 1846
DB 1649 ATQGDVLVLRNLTKVYRGORMPAVDRCLGIPPGECFGLLVNGAGKSTSTFRMVTGDTLA 1708
QY 1847 SRGSAVLAGHSAVAREPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVOTA 1906
DB 1709 SRGSAVLAGHSAVAREPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVOTA 1768

QY 1907 GSGIARLGSLWYADRPAGTSGGNKRKLATLALVGPVAVFLDEPTTGMDSARRFLWN 1966
DB 1769 GSGIARLGSLWYADRPAGTSGGNKRKLATLALVGPVAVFLDEPTTGMDSARRFLWN 1828
QY 1967 SLLAVREGSRVMTLSHMECEALCSRLAIWNGRFRCLGSPHLLKGRFAAGHTLTLRV 2026
DB 1829 SLLAVREGSRVMTLSHMECEALCSRLAIWNGRFRCLGSPHLLKGRFAAGHTLTLRV 1876
QY 2027 PAARSQPAAPFAAEFFGSELREAHGGLRLPQLPPGRCALARVFGELAVHGAHGVEDF 2086
DB 1877 ---RSQPAAPFAAEFFGSELREAHGGLRLPQLPPGRCALARVFGELAVHGAHGVEDF 1933
QY 2087 SVSQTMLLEEVFLYSKQKQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2146
DB 1934 SVSQTMLLEEVFLYSKQKQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 1993

RESULT 11
AAU04484
ID AAU04484 standard; protein; 1873 AA.
XX AC AAU04484;
XX DT 26-SEP-2001 (first entry)
XX DE Human PD-ATP-binding cassette (PD-ABC) protein form #2.
XX KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.
XX OS Homo sapiens.
XX PN WO200153490-AI.
XX PD 26-JUL-2001.
XX PF 23-JAN-2001; 2001WO-US002191.
XX PR 24-JAN-2000; 2000US-0177889P.
XX PR 30-JUN-2000; 2000US-0215405P.
XX PA (WARN) WARNER LAMBERT CO.
XX PI Johns MA, Tafuri SR, Wang M;
XX DR WPI; 2001-442259/47.
XX DR N-PSDB; AAS08707.
XX PT New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
XX of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.
XX PS Claim 10; Page 64-72; 77pp; English.
XX CC The sequence represents human PD-ATP-binding cassette (PD-ABC) protein
CC form 2. PD-ABC maps to chromosome 19p13.3 and is expressed in various
CC tissues including spleen, thymus, peripheral blood leukocytes, bone
CC marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to
CC diagnose and treat cardiovascular disorders, inflammatory disorders,
CC dyslipidaemia, epilepsy, diseases related to abnormal calcium flux,
CC coronary artery disease, Tangier's disease, familial high-density
CC lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
CC insulin resistance, obesity, alcoholism, retinal degeneration,
CC hypertension and vascular disease. The sequences are also used in drug
CC screening assays
XX Sequence 1873 AA;

Query Match 84.9%; Score 9460.5; DB 4; Length 1873;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1835; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 MAFWTQMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHHPLEHHECHFFNKP 60
DB 1 MAFWTQMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHHPLEHHECHFFNKP 60
QY 61 LPAGTVPWMLQGLICNNVNTCFPOLTPGEBPGRLSNFNDSLVSRLLADARTVLGGASAH 120
DB 61 LPAGTVPWMLQGLICNNVNTCFPOLTPGEBPGRLSNFNDSLVSRLLADARTVLGGASAH 120
QY 121 TLAGLGKLIATLAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGIALGQAQEP 180
DB 121 TLAGLGKLIATLAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGIALGQAQEP 180
QY 181 HSLEAAEDLAQELLALRSLVELRALIQRPRGTSGPLELSEALCSVRGFSSTVGP 240
DB 181 HSLEAAEDLAQELLALRSLVELRALIQRPRGTSGPLELSEALCSVRGFSSTVGP 240
QY 241 YEASDLMELVGPBPESALPDSSISLPACSELIGALDSHPILSRLLWRRUKPLILGKL 300
DB 241 YEASDLMELVGPBPESALPDSSISLPACSELIGALDSHPILSRLLWRRUKPLILGKL 300
QY 301 TPETKRLMAOVNRTFEELTLRDVREVMELGPRIFTFMNDSSNVAMQLQLQMDGRR 360
DB 301 TPETKRLMAOVNRTFEELTLRDVREVMELGPRIFTFMNDSSNVAMQLQLQMDGRR 360
QY 361 QPRPGGRDHMEALRSFLDPGSGGYSMQDAHADVGHVLTGLRVTETCLSLDKLEAPSEAA 420
DB 361 QPRPGGRDHMEALRSFLDPGSGGYSMQDAHADVGHVLTGLRVTETCLSLDKLEAPSEAA 420
QY 421 LVSRALQLLAHFRWAGVFLGPEDSDPTHEPTDLPDGPCHVIRIKRMDIDVTRTKIR 480
DB 421 LVSRALQLLAHFRWAGVFLGPEDSDPTHEPTDLPDGPCHVIRIKRMDIDVTRTKIR 480
QY 481 DRFWDGPAADPLTDLRYVMGGFYVLQDLVERAAVRVLSGANPRAGLYLQMPYPCVDD 540
DB 481 DRFWDGPAADPLTDLRYVMGGFYVLQDLVERAAVRVLSGANPRAGLYLQMPYPCVDD 540
QY 541 VFLRVSRLPLFLTLAWIYSVTLTKAVVREKETRLRDMRAMGLSRAVLWGLWFLSCL 600
DB 541 VFLRVSRLPLFLTLAWIYSVTLTKAVVREKETRLRDMRAMGLSRAVLWGLWFLSCL 600
QY 601 GPFLLSAAILLVILKGDILPYSHPGVFLFLAAFAVATVTSQFLLSAFSTRANLAAACG 660
DB 601 GPFLLSAAILLVILKGDILPYSHPGVFLFLAAFAVATVTSQFLLSAFSTRANLAAACG 660
QY 661 GLAYFSLYLPVLCVAVWRDLRDPAGRVAASLLSPVAFGFCESLALLEEGEGAQWNVG 720
DB 661 GLAYFSLYLPVLCVAVWRDLRDPAGRVAASLLSPVAFGFCESLALLEEGEGAQWNVG 720
QY 721 TRPTADVFLSAQVSGLLLDAAALYGLATWYLEAVCPQYGIPEPWNFPFRSSWCGPRPP 780
DB 721 TRPTADVFLSAQVSGLLLDAAALYGLATWYLEAVCPQYGIPEPWNFPFRSSWCGPRPP 780
QY 781 KSPAPCTPDPKVLVEEAPPGLSPGVSVSRLEKRPFGSPQALRGSLDFYQGHITAF 840
DB 781 KSPAPCTPDPKVLVEEAPPGLSPGVSVSRLEKRPFGSPQALRGSLDFYQGHITAF 840
QY 841 GHNGAGKTTTLLSILSGIFPPSGGSAFTLGHDRVSSMAAIPHLGVCQYNNVLFDM 900
DB 841 GHNGAGKTTTLLSILSGIFPPSGGSAFTLGHDRVSSMAAIPHLGVCQYNNVLFDM 900
QY 901 HVMFYGRKGLSAAVVGPEODRLLDVGLSVKQSVQTRHLSGGQKRLSVAIAFVGSQV 960
DB 901 HVMFYGRKGLSAAVVGPEODRLLDVGLSVKQSVQTRHLSGGQKRLSVAIAFVGSQV 960
QY 961 VILDEPTAGVDPASRRGIIWELLILKYREGRTLILSTHLLDEAELLGDRVAVAGRLCCG 1020
DB 961 VILDEPTAGVDPASRRGIIWELLILKYREGRTLILSTHLLDEAELLGDRVAVAGRLCCG 1020
QY 1021 SPLFLRRHLGSGVYTLVKARLPLTTNEKADTMEGSDVTQBEKNGSQGSRVGT 1080

Db 1021 SPFLRLHSGSYLTNFKARLPLTTNEXADTDMGSDVTRQKNGSQSGVGTQPLIA 1080
Qy 1081 LVQHWVPGARLVEELPHELVVLVPTGADHGSFATLFRDLRLAELRLTGVGISTSL 1140
Db 1081 LVQHWVPGARLVEELPHELVVLVPTGADHGSFATLFRDLRLAELRLTGVGISTSL 1140
Qy 1141 EFLKVBECADTDMEDSCQCHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
Db 1141 EFLKVBECADTDMEDSCQCHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
Qy 1201 QSGSDPAVGRVQGWALTQQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFLIV 1260
Db 1201 QSGSDPAVGRVQGWALTQQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFLIV 1260
Qy 1261 PPFHYPALRLSPTWYGAQVSPFSDAPGDPGRARLLEALLQEAELGEPVVOHSHRFA 1320
Db 1261 PPFHYPALRLSPTWYGAQVSPFSDAPGDPGRARLLEALLQEAELGEPVVOHSHRFA 1320
Qy 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTSGSEVVQ 1380
Db 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTSGSEVVQ 1380
Qy 1381 NLTGRLSDFLVKTYPRLVRQGLKTKWVNEVRYGFSGLGRDPGLPGQELGRSVEELW 1440
Db 1381 NLTGRLSDFLVKTYPRLVRQGLKTKWVNEVRYGFSGLGRDPGLPGQELGRSVEELW 1440
Qy 1441 ALLSPLPGALDRVLKNTTAWAHSJDAQDSLKIFWNNKGWHSWAFVNRASNAIIRAHLP 1500
Db 1441 ALLSPLPGALDRVLKNTTAWAHSJDAQDSLKIFWNNKGWHSWAFVNRASNAIIRAHLP 1500
Qy 1501 PGPARHAHSITTLNHPNLTKQLSEALMASSVDVLSICVVFAMSFVPASFTLVLE 1560
Db 1501 PGPARHAHSITTLNHPNLTKQLSEALMASSVDVLSICVVFAMSFVPASFTLVLE 1560
Qy 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNVLPACIVLVIFLAFQQRAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNVLPACIVLVIFLAFQQRAYVAPANLPALL 1620
Qy 1621 LLLLYGWSITPLMYPASFPFSPVSTAYVVLTCINLFTGNGSMATFVLEFSDOKLOEV 1680
Db 1621 LLLLYGWSITPLMYPASFPFSPVSTAYVVLTCINLFTGNGSMATFVLEFSDOKLOEV 1680
Qy 1681 SRLKQVFLIPHFCLGRGLDMVNRQAMADAFERLGRDQFQSPLRWEVVGKLLAMVIQ 1740
Db 1681 SRLKQVFLIPHFCLGRGLDMVNRQAMADAFERLGRDQFQSPLRWEVVGKLLAMVIQ 1740
Qy 1741 GPLFLFTLLQHRSQLPQPRVRSPLILGEBEDVARERERVQATGQDVVLRLNLT 1800
Db 1741 GPLFLFTLLQHRSQLPQPRVRSPLILGEBEDVARERERVQATGQDVVLRLNLT 1800
Qy 1801 VYRGORMPAVDRCLGIPGECFGLLGVNGAGKTSTFRMVTGDTLASGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRCLGIPGECFGLLGVNGAGKTSTFRMVTGDTLASGEAVLAGHSVAR 1860
Qy 1861 EPSAHLMSG-YCQSDAIFELL 1882
Db 1855 -----HLLGTYCM---PIFVLL 1869

RESULT 12

ABG72695
ID ABG72695 standard; protein; 2167 AA.
XX
AC ABG72695;
XX
DT 10-MAR-2003 (first entry)
XX
DE Mouse ATP-binding cassette transporter-like protein, ABCL.
XX
KW Mouse; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;

nervous system disorder; Stargardt disease; degenerative disorder;
inflammatory retinopathy; cystic fibrosis; multidrug resistance;
lymphoid condition; myeloid cell condition; AIDS; lymphoma;
acquired immunodeficiency disorder; leukaemia; neutropenia; anaemia;
autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
hypothalamus disorder; obesity; diabetes; reproductive disorder;
energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
autoimmune disease; inflammatory disease; multiple sclerosis.

Mouse musculus.

Key Location/Qualifiers
Peptide 1..46
Protein /label= Signal_peptide
47..2167
/label= Mature_ABCL

US2002127647-A1.

12-SEP-2002.

28-NOV-2001; 2001US-00995542.

28-NOV-2000; 2000US-0253520P.

(SHUT/) SHUTTER J.
(ULIA/) ULIAS L.

Shutter J, Ulia L;

WPI; 2003-147394/14.

N-PSDB; ABX14665.

Novel ATP-binding cassette transporter-like polypeptides and

polynucleotides useful for diagnosing, preventing, treating disorders
involving immune, nervous system, thyroid, hypothalamus and impaired
transport of lipids.

Claim 13; Fig 1; 149pp; English.

The invention relates to an isolated murine and human ATP-binding
cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
3111. Also include are the nucleic acids encoding the ABCL proteins,
vectors, host cells, ABCL binding agents, a selective binding agent or
its fragment comprising at least one complementarity determining region
(CDR) with specificity for ABCL which (produced by immunising an animal
with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
fusion polypeptide, a device comprising a membrane suitable for
implantation (permeable to the protein and impermeable to materials
derivative to the cells, and cells encapsulated within the membrane)
where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
and modulators are useful for the diagnosis and/or treatment of diseases
and conditions involving impaired transport of lipids, including
cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
involving functional and trophic disturbances of the nervous system such
as Stargardt disease, degenerative and inflammatory retinopathy, cystic
fibrosis, conditions involving multidrug resistance, conditions involving
lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
neutropenia, anaemia and autoimmune diseases, conditions involving the
thyroid e.g. hyper and hypothyroidism; conditions involving the
hypothalamus including obesity, diabetes, reproductive disorders, energy
balance disorders, peripheral neuropathies including myelinopathies and
axonopathies, autoimmune and inflammatory diseases involving the nervous
system including multiple sclerosis. The present sequence represents
murine ABCL

Sequence 2167 AA;

Query Match 77.3%; Score 8613.5; DB 6; Length 2167;
Best Local Similarity 76.7%; Pred. No. 0;

Matches 1678; Conservative 163; Mismatches 283; Indels 65; Gaps 11;

```
Qy 1 MAFWTQMLLWKQNFMTARRQPVOLLVELLWPLFLFVLVAVRSHHPLEHHECHFFNKP 60
Db 1 MAFCTQLMLLWKNYTYRRQPIQLLVELLWPLFLFVLVAVRSHHPLEHHECHFFNKP 60
Qy 61 LPSAGTVPWLGGLICNVNNTCFPOLTPCEEPGRLSNFENDSLVSRLLADARTVLVGGASAH 120
Db 61 LPSAGTVPWLGGLICNVNNSCFQHPPTCEKPGVLSNFKDSLISRLADARTVLVGGHIO 120
Qy 121 TLAGLGLIATRAARSTAQP----QPTKQSPLEPPMLDVAELTSLLRTESTLGLALGOA 176
Db 121 MLDLXGLIPVLRVAGGARQESDQPTSQ-----SVTKLEKILQASLDPLVGOA 173
Qy 177 QEPHLSLEAAEDLAQELLALRSIVELRALLQRPRGTSGPLLELSEALCSVRGSSVGP 236
Db 174 QDSMRKFSDAIRDLAQELLTLPSIMELRALLRRPRGSAGSLSELVSEALCSTKGPSSPG 233
Qy 237 SLNWVEASDLWELVGOBPESALPDSSISPACSELIIGALDHPGLSRLLWRRLLKPLILG 296
Db 234 SLNWYEANQLNEFMGPVAPALPONSISPACSEFVGLDDHPVSRLLWRRLLKPLILG 293
Qy 297 FAPDTPTRKLMAQVNRTEFELTLRLDVRVWEMLGPRIFFTMNDSNVAMQLRLQMD 356
Db 294 FAPDTPTRKLMAQVNTFEELALRLDLHELWGLVGPQIFNMNDSNVAMQLRLQMD 353
Qy 357 EGRQPPRGDRHWEALRSFLDPSGGYSWODAHADVGLVGTIGRVTESLSDKLEAAP 416
Db 354 TGORQOTPRAQKLEAIKDFLDPSRGGSYWRHADMGRLAGILGQWMECVSLDKLEAVP 413
Qy 417 SEALVSRALQLLAERHFWAGVFLGPDSSDPTTEHPTDLPDGPCHVKIRMDIDVVRT 476
Db 414 SEALVSRALQLGERLLWAGIVFLSPHPLDPSBELSPALSPGHLRFKIRMDIDVVRT 473
Qy 477 NKIRDREWDPGPAADPLTLDRVYWGFFVYLQDLVERAAVRLSGANPRAGILYLOQMPYC 536
Db 474 NKIRDKFWDPGCSADPFMDLRYVWGGFVYLQDLLEQAARVVLGGNSRTGLYLOQMPHC 533
Qy 537 YVDVDFLRVLSRSLPLFLTLAWISVTLTVKAVVREKETRLDTRMANGLSRAVLWLGWF 596
Db 534 YVDVDFLRVLSRSLPLFLTLAWISVTLTVKAVVREKETRLDTRMANGLSRAVLWLGWF 593
Qy 597 LSCGLGFLLSAALAVLVLKGDILPYSHPGVVFLPAAFAVATVTSQFLLSAFPSRANLA 656
Db 594 LSCGLGFLVLSAALLVLVLKGNILPYSHPVVIFLAAFAVATVVAQFLLSAFPSRANLA 653
Qy 657 AACGGLAYFSLPYLVLCVNRDLRDPAGGRVAASLLSPVAFGFCESLALLEEGEGAQW 716
Db 654 AACGGLAYFALYPVLCVNRRLHLGLLAAALLSPVAFGFCESLALLEEGDGAQW 713
Qy 717 HNVGTRPTADVFSIAQVSGLLLLDAALYGLATWYLEAVCPGQYGIPEPWNPPFRSTWCG 776
Db 714 HNLGTGPAEDVFSIAQVSAFLLLDVAVIYGLALWYLEAVCPGQYGIPEPWNPPFRSTWCG 773
Qy 777 PRPKSPAPCTPLDPKVLVEEAPGLSPGVSVESLEKRPGPSQPALRGISLDPYOGHI 836
Db 774 PGPKSVLAPAPQDPKVLVEEPGLVPVGSIRGLKKHFGCCQPALQGLNLDIFYEGHI 833
Qy 837 TAFLGHNGAGKTTTSLISGLFPSPSGSASFILGHDVRSMAAIRPHLGVCPCQYNVFDML 896
Db 834 TAFLGHNGAGKTTTSLISGLFPSPSGSASILGHDVQTNMAAIRPHLGICPCQYNVFDML 893
Qy 897 TVDEHVPFYGRKLGLSAAVGPEDODRLDVLVSKQSVQTRHLSGGMQRKLSVAIAFVG 956
Db 894 TVEEHVPFYGRKLGVSAAMGPPEERLIRDVGLTKRDTQTRHLSGGMQRKLSVAIAFVG 953
Qy 957 GSQVVLDEPTAGVDPASRRGIWELLKYRREGRTLILSTHLLDEAELLGDRVAVVAGRL 1016
Db 954 GSRVIVDEPTAGVDPASRRGIWELLKYRREGRTLILSTHLLDEAELLGDRVAVVAGSL 1013
Qy 1017 CCCGSPFLRRHLHSGGYTLVVKARLPLTTNE--KADTDMEGSVDTROKKNQSGS---- 1071
Db 1014 CCCGSPFLRRHLHSGGYTLVVKSSQSLVTHDAKGDSE----DPRREKKS DGNRTSDT 1068
```

```
Qy 1072 -----RVGTPOLLALVOHWVPGARLVEELPHELVLVLPTTGAHD 1110
Db 1069 AFTRGTSKSNQAPAGVAPITPSTARILELVOHQVPGAOLVEDLPHELHLLVLYAGALD 1128
Qy 1111 GSFATIFRELDTRLAELRTGYGISTDSLEEEIFLKVVEECAADTDMEDSGCGHLCGTGIA 1170
Db 1129 GSFAMVFOELDQOELLELGLTGYGISTDNLEEEIFLKVED--AHREGGDSRQPLHRT--- 1183
Qy 1171 GLDVTURLKWPPOETALENGEPAGSA--PETDQSGSDPAVGRVOGWALTQQOALLKLR 1228
Db 1184 ---CTQPPTPGPPEASVLENGELAKLVLDPOAPQGLAPNA-AQVGQWTLTCQQLRALLH 1239
Qy 1229 FILLARRSRGIFAQIVLPALFVGLALVPSLIVPBGHYPALRLSPTWYGAQVSPFSSEDAP 1288
Db 1240 FLLARRSRGIFAQIVLPALFVGLALVPSLIVPBGHYPALRLSPTWYGAQVSPFSSEDAP 1299
Qy 1289 GDFCARLLEALLQEALEBPPVQHSR-----FSAPVPAEPAEVAKVILASGNWT 1337
Db 1300 GDFNEMKLEALLGEAGLQEPSMQDKARGSECTHSLACYFTVPEVPPDVASILASGNWT 1359
Qy 1338 PESPSACQSQCGARRLLPDCPAAAGPPPPPOAVTQSGEVVQNLTORNISDFILVKTYPR 1397
Db 1360 PESPSACQSQCGARRLLPDCPAGAGPPPPPOAVAGLGEVQNLTCGRNVSDFLVKTYPS 1419
Qy 1398 LVROGLTKKVMNEVRYGGFSLGGRDPLPSGOELGSESVEELWALLSPLPGALDRVLKN 1457
Db 1420 LVRRGLTKKVMNEVRYGGFSLGGRDPLPTGHEVVKTLAEIRALLSPQGNADRLINN 1479
Qy 1458 LTAWAHSLDAQDSLIKIFWNNKGWHSMAVFNVRASNAILRAHLPPGPAPARHAHSITTLN 1517
Db 1480 LTQWALGLDARNSLIKIFWNNKGWMAVFNVRANNGLHALLPSGVPVRAHHSITTLN 1539
Qy 1518 NLTKQELSEALMASSVDVLVSCVVPFAMSFVPASFTVLVLIBERVTRAKHLQMLGGLSPT 1577
Db 1540 NLTKQELSEALMASSVDVLVSCVVPFAMSFVPASFTVLVLIBERVTRAKHLQMLGGLSPT 1599
Qy 1578 LYWLGNFLWDMCNVLPACIVVLIFLAFQORAVVAPANLPAALLLALLLYGWSITPLMYP 1637
Db 1600 LYWLGNFLWDMCNVLPACIVVLIFLAFQORAVVAPANLPAALLLALLLYGWSITPLMYP 1659
Qy 1638 SFFPSVPSTAYVLTCLNLFINGMSMATFVLELFDQKLQEVSRILKQVFLPPHFCGLG 1697
Db 1660 SFFPSVPSTAYVLTCLNLFINGMSMATFVLELSDQNLQEVSRILKQVFLPPHFCGLG 1719
Qy 1698 RGLIDWVRNQAMADAFERLGDQFQSPLRWVGVKNLLAMVIOGPIFLTLTLLQHRSQL 1757
Db 1720 RGLIDWVRNQAMADAFERLGDQFQSPLRWVGVKNLLAMVIOGPIFLTLTLLQHRSQL 1779
Qy 1758 LPQPRVRSPLPCEDEEDVARERERVVGATQGDVLRNLTKVYRGORMPAVDRLCLGI 1817
Db 1780 LPQSKRLLPPLCEDEEDVARERERVVGATQGDVLRNLTKVYRGORMPAVDRLCLGI 1839
Qy 1818 PPCECFGLLVGNGAGTKSTFRMVTGDTLARGAVLAGHSVAREPSAAHLSMGYCPQSDA 1877
Db 1840 PPCECFGLLVGNGAGTKSTFRMVTGDTLPSGGEAVLAGHNAQERSAAHLSMGYCPQSDA 1899
Qy 1878 IFELLTGREHLELLARLURGVPEAQVATAGSGLARIGLSWYADRPAGTYSGGNKRKLATA 1937
Db 1900 IFDILLTGREHLELLARLURGVPEAQVATAGSGLARIGLSWYADRPAGTYSGGNKRKLATA 1959
Qy 1938 LALVGDPVAVFLDEPTTGMDPSARRFLWNLSILAVVREGSVMLTSHSMECEALCSRLAI 1997
Db 1960 LALVGDPVAVFLDEPTTGMDPSARRFLWNLSILAVVREGSVMLTSHSMECEALCSRLAI 2019
Qy 1998 MVNGRFRCLSGSPQHLKGRFAAGHTLTLRVPAARSQPAFAAFVAFPGSELREAHGGRLRF 2057
Db 2020 MVNGRFRCLSGSPQHLKGRFCAGHTLTLRVPPDQPEPAIFRITFPCAELREVHGSRLRF 2079
Qy 2058 QLPFGGRCALARVFGELAVHGAHGVDEDFSVQSTMLEEVFLYFSKQDGKDEDEEQKEAG 2117
Db 2080 QLPFGGRCTLTRVPRELAAQAGRAHGVDFSVQSTMLEEVFLYFSKQDGKEESSRQ-EAE 2138
```

QY 2118 VGVDPAPGLQHPKRVSQFLDDPSTARTVL 2146
DB 2139 BEEVSKPGRQPKRVSRFLEDPSSVETMI 2167

RESULT 13
ABG72697
ID ABG72697 standard; protein; 1550 AA.
AC ABG72697;
XX
XX 10-MAR-2003 (first entry)
XX
DE Human ATP-binding cassette transporter-like protein, ABCI1550.
XX
KW Human; ATP-binding cassette transporter-like protein; ABCI1550;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX
OS Homo sapiens.
XX
XX US2002127647-A1.
XX
XX 12-SEP-2002.
XX
XX 28-NOV-2001; 2001US-00995542.
XX
XX 28-NOV-2000; 2000US-0253520P.
XX
XX (SHUT/) SHUTTER J.
XX (ULIA/) ULIAS L.
XX
XX Shutter J, Ulias L;
XX
XX WPI; 2003-147394/14.
XX N-PSDB; ABX14667.
XX
XX Novel ATP-binding cassette transporter-like polypeptides and
XX polynucleotides useful for diagnosing, preventing, treating disorders
XX involving immune, nervous system, thyroid, hypothalamus and impaired
XX transport of lipids.
XX
XX Claim 13; Fig 3; 149pp; English.

The invention relates to an isolated murine and human ATP-binding cassette transporter-like (ABCL) polypeptide, or the amino acid sequence encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-3111. Also included are the nucleic acids encoding the ABCL proteins, vectors, host cells, ABCL binding agents, a selective binding agent or its fragment comprising at least one complementarity determining region (CDR) with specificity for ABCL which (produced by immunising an animal with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL fusion polypeptide, a device comprising a membrane suitable for implantation (permeable to the protein and impermeable to materials detrimental to the cells, and cells encapsulated within the membrane) where the cells secrete ABCL, an ABCL transgenic non-human mammal and an array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids and modulators are useful for the diagnosis and/or treatment of diseases and conditions involving impaired transport of lipids, including cardiovascular disease, hypertriglyceridaemia, atherosclerosis, hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions involving functional and trophic disturbances of the nervous system such as Stargardt disease, degenerative and inflammatory retinopathy, cystic fibrosis, conditions involving multidrug resistance, conditions involving lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,

CC neutropaenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies, autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL truncated variant, ABCI1550
XX
SQ Sequence 1550 AA;

Query Match 68.2%; Score 7596; DB 6; Length 1550;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 167 ESLGALGAQOEPLHSLLEAAEDLAQELLALSVELRALLOPRGTSGLLESLALCS 236
DB 29 ESLGALGAQOEPLHSLLEAAEDLAQELLALSVELRALLOPRGTSGLLESLALCS 88
QY 227 VRGPESTVGPSSLNWYEASDLMELVQOEPEPESALPDSLSLSPACSELIGALDHPHLLMRR 286
DB 89 VRGPESTVGPSSLNWYEASDLMELVQOEPEPESALPDSLSLSPACSELIGALDHPHLLMRR 148
QY 287 LKPLILGKLLFAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIFTFMDSSNVA 346
DB 149 LKPLILGKLLFAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIFTFMDSSNVA 208
QY 347 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTSC 406
DB 209 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTSC 268
QY 407 LSLDKLEAAPSEAAALVSRALQLLAEHFRWAGVVFGLPEDSSDPTHEPTDGLPGHVRKI 466
DB 269 LSLDKLEAAPSEAAALVSRALQLLAEHFRWAGVVFGLPEDSSDPTHEPTDGLPGHVRKI 328
QY 467 RMDIDVTRTKIRDRFWDGPAADPLTDLRYWVGGFVYLQDLVERAAVRLSGANPRAG 526
DB 329 RMDIDVTRTKIRDRFWDGPAADPLTDLRYWVGGFVYLQDLVERAAVRLSGANPRAG 388
QY 527 LYLOQMPYPCVVDVFLRVLSRLPFLTLAWIYSVTLTVKAVVREKETRLDTRWAMGL 586
DB 389 LYLOQMPYPCVVDVFLRVLSRLPFLTLAWIYSVTLTVKAVVREKETRLDTRWAMGL 448
QY 587 SRAVLWLGWFLSCGLPFLLSAALLVLKLGDLIPYSHPGVVFLLAFAVATVTSFLL 646
DB 449 SRAVLWLGWFLSCGLPFLLSAALLVLKLGDLIPYSHPGVVFLLAFAVATVTSFLL 508
QY 647 SAFTSRANLAAACGLAYFSLYPLVCVAVNRDRLPAGGRVAAALLSPVAFGFCESLAL 706
DB 509 SAFTSRANLAAACGLAYFSLYPLVCVAVNRDRLPAGGRVAAALLSPVAFGFCESLAL 568
QY 707 LEEQEGGAOMHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGQYGIPEPNW 766
DB 569 LEEQEGGAOMHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGQYGIPEPNW 628
QY 767 FPFRRSYWCGPRPKSPACPTPLDPAKVLVEEAPGLSPGVSVRSLEKRFPGSPQALRG 826
DB 629 FPFRRSYWCGPRPKSPACPTPLDPAKVLVEEAPGLSPGVSVRSLEKRFPGSPQALRG 688
QY 827 LSLDFYQGHITAFIAGHNGAGTKTTLSILSGLPFPPSGGSAFILGHVDRSSMAAIRPHLGV 886
DB 689 LSLDFYQGHITAFIAGHNGAGTKTTLSILSGLPFPPSGGSAFILGHVDRSSMAAIRPHLGV 748
QY 887 PQYNVLFDMLTVDHFWFVYGRKLGISAADVPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 946
DB 749 PQYNVLFDMLTVDHFWFVYGRKLGISAADVPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 808
QY 947 KLSVAIAFVGGSVQVILDEPTAGVDPASRRGIWELLKYREGRTILSTHHLDEALLGD 1006
DB 809 KLSVAIAFVGGSVQVILDEPTAGVDPASRRGIWELLKYREGRTILSTHHLDEALLGD 868
QY 1007 RVAVVAGRLCCCGSPLFLRHLSGGYYTLVVKARLPLTTNEKADTMEGSVDTTQEKKN 1066
DB 869 RVAVVAGRLCCCGSPLFLRHLSGGYYTLVVKARLPLTTNEKADTMEGSVDTTQEKKN 928

```
QY 1067 GSQSRVGTPLLALVOHWVPGARLVPELPHLVLPYTGAGHDSFATLFRELDTRLAE 1126
DQ 929 GSQSRVGTPLLALVOHWVPGARLVPELPHLVLPYTGAGHDSFATLFRELDTRLAE 988
QY 1127 LRLTGYGISTSLBEIFLKVVVECAADTDMEDGSCGOHLCTGTAGLDVTLRLKMPPOETA 1186
DQ 989 LRLTGYGISTSLBEIFLKVVVECAADTDMEDGSCGOHLCTGTAGLDVTLRLKMPPOETA 1048
QY 1187 LENEPAGSPETDQSGPDAGVRVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1246
DQ 1049 LENEPAGSPETDQSGPDAGVRVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1108
QY 1247 ALFVGLALVSLIIPPFGHPALRLSPMTYGAQVSFESEADPGDGRARLLEALLQEAGL 1306
DQ 1109 ALFVGLALVSLIIPPFGHPALRLSPMTYGAQVSFESEADPGDGRARLLEALLQEAGL 1168
QY 1307 EEPVQVHSSHRFSAPEVAEVAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGGP 1366
DQ 1169 EEPVQVHSSHRFSAPEVAEVAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGGP 1228
QY 1367 PPQAVTGSVEVQNLGRNLSDFLVTPRLVRQGLTKKWNVEVRYGFSLGGRDPGL 1426
DQ 1229 PPQAVTGSVEVQNLGRNLSDFLVTPRLVRQGLTKKWNVEVRYGFSLGGRDPGL 1288
QY 1427 PSGOELGRSVEELWALLSPGGLDRVLKNTAWAHSLSLDAQDSLKTIWNNKGWHSVAF 1486
DQ 1289 PSGOELGRSVEELWALLSPGGLDRVLKNTAWAHSLSLDAQDSLKTIWNNKGWHSVAF 1348
QY 1487 VNRSNAILRAHLPFGPARHAHSITTLNHPNLTKQLESAALMASSVDVLVSICVVFAM 1546
DQ 1349 VNRSNAILRAHLPFGPARHAHSITTLNHPNLTKQLESAALMASSVDVLVSICVVFAM 1408
QY 1547 SFVPASTVLIERVTRAKHLQMGSLPTLYWGNFLMDMCMNYLPACIVLILFAQ 1606
DQ 1409 SFVPASTVLIERVTRAKHLQMGSLPTLYWGNFLMDMCMNYLPACIVLILFAQ 1468
QY 1607 QRAYVAPANLPALLLLLLLVCWSITPL 1633
DQ 1469 QRAYVAPANLPALLLLLLLVCWSITPL 1495
RESULT 14
AAB38111
ID AAB38111 standard; protein; 2261 AA.
AC AAB38111;
XX XX
DT 29-JAN-2001 (first entry)
XX XX
DE Human ABC1 cholesterol transporter mutant, V771M.
XX XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW x-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutin.
XX XX
OS Homo sapiens.
XX XX
PN WO200055318-A2.
XX XX
PD 21-SEP-2000.
XX XX
PF 15-MAR-2000; 2000WO-IB0000532.
XX XX
PR 15-MAR-1999; 99US-0124702P.
PR 08-JUN-1999; 99US-0138048P.
PR 17-JUN-1999; 99US-0139600P.
```

```
PR 01-SEP-1999; 99US-0151977P.
XX XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX XX
XX Hayden MR, Wilson AR, Pimstone SN;
XX WPI; 2000-587528/55.
XX XX
XX New ABC1 polypeptide is useful for treating diseases associated with ABC1
XX biological activity, e.g. Alzheimer's disease, Huntington's disease and
XX cancer.
XX XX
XX Example; Page; 229pp; English.
XX XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds. It
CC further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or
CC prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, x-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: A7012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease. Note: The present sequence is not shown
CC in the specification, but is derived from the native human ABC1 shown on
CC pages 152-157
XX XX
SQ Sequence 2261 AA;
Query Match 51.8%; Score 5769.5; DB 3; Length 2261;
Best Local Similarity 50.3%; Pred.No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;
QY 1 MAFVQLMLLWKNFMVRRRQPVQLLVLLWLPFLFFILVAVRHSHPLEHCHFFNKP 60
DQ 1 MACWPQLRLLLWKNLTFRRQTCLLLEAVAPLFIILISVRLSYPPYQHECHFFNKA 60
QY 61 LPSAGTVPLWQLGILCNVNNTCFPOLTPGEBFGRISNFNDSLVSRLLADARTVLGASAH 120
DQ 61 MPSAGTLPWVQGIICNANNPCFRYPTGEPAGVGVGNFNKSIARLFSRARLLYSOKDT 120
QY 121 TLAGLGLIATLRAARSTAQ-----PQTKQSPLEPPML-- 154
DQ 121 SMKDMRKVLRTLQOIKKSSNKLKQDFLVNNETPSGLFVNNLSLPKSTVDKMLRADVLH 180
QY 155 -----DVAEL-----LTSLL 164
```


CC	AAB38114	standard; protein; 2261 AA.
CC	AAB38114;	
CC	29-JAN-2001	(first entry)
CC	Human ABC1 cholesterol transporter mutant, E1172D.	
CC	Human ABC1 cholesterol transporter; chromosome 9q31;	
CC	ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;	
CC	Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;	
CC	cardiovascular disease; coronary artery disease; coronary restenosis;	
CC	cardiovascular disease; peripheral vascular disease;	
CC	Alzheimer's disease; Niemann-Pick disease; Huntington's disease;	
CC	X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;	
CC	prognosis; prophylaxis; drug screening; transgenic animal; mutant;	
CC	Mutein.	
CC	Homo sapiens.	
CC	W0200055318-A2.	
CC	21-SEP-2000.	
CC	15-MAR-2000; 2000WO-IB0000532.	
CC	15-MAR-1999; 99US-0124702P.	
CC	08-JUN-1999; 99US-0138048P.	
CC	17-JUN-1999; 99US-0139600P.	
CC	01-SEP-1999; 99US-0151977P.	
CC	(UTBR-) UNIV BRITISH COLUMBIA.	
CC	(XENO-) XENON BIORESEARCH INC.	
CC	Hayden MR, Wilson AR, Pimstone SN;	
CC	WPI; 2000-587528/55.	
CC	New ABC1 polypeptide is useful for treating diseases associated with ABC1	
CC	biological activity, e.g. Alzheimer's disease, Huntington's disease and	
CC	cancer.	
CC	Example; Page; 229pp; English.	
CC	The invention relates to the human ABC1 cholesterol transporter protein	
CC	(B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is	
CC	a member of the ATP-binding cassette (ABC transporter) superfamily of	
CC	proteins, and plays a crucial role in cholesterol transport, particularly	
CC	intracellular cholesterol trafficking in monocytes and fibroblasts, being	
CC	involved in cholesterol efflux from the cell. The gene encoding ABC1 is	
CC	located on chromosome 9q31, and mutations in this gene are associated	
CC	with two genetic HDL (high density lipoprotein) deficiency disorders,	
CC	Tangier disease (TD) and familial HDL deficiency (FHA). These diseases	
CC	are distinguishable in that TD is an autosomal recessive disorder, while	
CC	FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good	
CC	cholesterol") in the blood correlate with a high risk of cardiovascular	
CC	disease, particularly coronary artery disease, but also cerebrovascular	
CC	disease, coronary restenosis, and peripheral vascular disease.	
CC	Conversely, a high level of HDL has protective effects against	
CC	cardiovascular disease. The invention provides genetic constructs and	
CC	transgenic cells and non-human animals comprising human ABC1 nucleic	
CC	acids, and methods of gene therapy for the treatment or prevention of	
CC	cardiovascular disease comprising the administration of an expression	
CC	vector encoding ABC1 or an active fragment thereof. The invention also	
CC	encompasses compounds which mimic ABC1 activity, compounds which	
CC	stimulate ABC1 expression and methods of screening for such compounds. It	
CC	further relates to methods for determining whether a patient has an	
CC	increased risk for cardiovascular disease due to polymorphisms in the	
CC	ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or	
CC	prevent cardiovascular disease, especially coronary artery disease,	
CC	cerebrovascular disease, coronary restenosis or peripheral vascular	
CC	disease. They may also be used in the treatment of diseases associated	
CC	with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick	
CC	disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.	
CC	The invention specifically excludes proteins with the exact amino acid	
CC	sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic	
CC	acid with the exact sequence as GenBank Accession No: AJ012376.1. The	
CC	present sequence represents a mutant human ABC1 cholesterol transporter	
CC	associated with an altered cholesterol level and therefore an altered	
CC	risk of cardiovascular disease. Note: The present sequence is not shown	
CC	in the specification, but is derived from the native human ABC1 shown on	
CC	pages 152-157	
CC	Query Match 51.8%; Score 5769.5; DB 3; Length 2261;	
CC	Best Local Similarity 50.3%; Pred. No. 0;	
CC	Matches 114; Conservative 368; Mismatches 606; Indels 161; Gaps 26;	
CC	1 MAFWQMLMLLWKNFMYRRRQPVQLLVLELLWPLFFILVAVRHSHPPLEHCHFPNKP 60	
CC	1 MACWPQLRLLWKNLTFRRRQTCQLLEVAWPLFIFILISVRLSPYPPYEQHECHFPNKA 60	
CC	61 LPSAGTVPMLOGLICNVNNTCPQLTFGEHPGLSNFNDLSVRLSLADARTVLGGASAH 120	
CC	61 MPSAGTLPWVGIGICNANNPFCFRYPTEGAPGVGVGNFNKSIVARLFSDAERLLYSQOT 120	
CC	121 TLAGLGKLIATRAARSTAQ-----POPTKQSPLEPPML-- 154	
CC	121 SMKMRKVLKTLQOIKKSSNLKQLDVLVDNFTSGFLYHNLSPKSTVDKMLRADVILH 180	
CC	155 -----DVAEL----- 164	
CC	181 KVFLQGVQLHLTSLCNGSKSEMIQLGDQVSELGLPREKLAARVLRSLNMDILKPI 240	
CC	165 RT--ESLGLALGOAQEPHLSLLEAAEDLAQELLALSLVELR---ALLQPRGTSGPLE 219	
CC	241 RTLNSTSPFPEKELAEATKTLHSLGTLAQELFSMRSDMRQVEMPLTNVNSSSTQI 300	
CC	220 ---LSEALCSVRGSPSTVGPGLNWEASDLMELVG---OEPESALPDSLSLSPACSELIG 272	
CC	301 YQAVSRIVCGHPGGGLKLSLNWYEDNNYKALFGNGTEEDAEFTVDNSTTTCYNDLMK 360	
CC	273 ALDHPHLSRLMLRRLKPLILGLKLLFAPDTPFTFKMAQVNRTPFEELTLDRVREWMELG 332	
CC	361 NLESSPLSRITIKALKPLLVGKILYTPDTPATQVMAEVNKTFOELAVFHDLEGMWELS 420	
CC	333 PRFTFMNDSSNVAMQLRLQLOMDEGR-RQPRGGRD-----HWEALRSFLDP 379	
CC	421 PKIWTFMENSQENDLVRMLLDSRDNDHFWEQQLDGLDWTQADIVAFLAKHPEDVQS--- 476	
CC	380 GSGG--YSWODAHADVGLVGTTLGRVTECLSLDKLEAAPSEAAALVSRALQLLAHREFWAG 437	
CC	477 -SNGSVYTWREAFNETNQAIRTISRPMECVNLKLEPIATEVWLINKSMELLDERKEFWAG 535	
CC	438 VFLGPEPDSDDPTEHPTDPLGPGHVRIRKIMDIDVVTNRKIRDRFWDPGPAADPLTDLR 497	
CC	536 IVFTGTPGSIELPH-----HVYKIRMDIDNVERTNKIKDGYWDGPRADPFEDMR 587	
CC	498 YWGGFVYLODLVERAARVLSGANPRAGLYLOMPYCVYDDVFLVLSRSPLFLTLA 557	
CC	588 YWGGFAYLODVVEQAIIRVLTGTEKTKYGYMOQMPYCYVDDIFLVMGRSPLFLMTLA 647	
CC	558 WYYSVTILTVKAVVREKETLRDTRMAGLSRAVLWGLWFLSLGPPPLLSAALVLVLKLG 617	
CC	648 WIYSVAVILIGIYVEKARUKETNRIMGLNSLWFSWFISSILPLVLSAGLLVILKLG 707	
CC	618 DILPYSHPGVVFLLFAFAVATVYQSFLLSAFFSRANLAAACGLAYFSYLYPLVYLCVAV 677	
CC	708 NLIPYSDPSVVFVFLSVFAVVTILQCFILSTLSRANLAAACGGIITYFTLYPLVYLCVAV 767	
CC	678 RDLRPPAGRVASILLSPVAFPGCESLALLEBOGEGQWNVGTRPT-ADVFSLAQVSLG 736	
CC	768 QDYVGTFTKIFASILLSPVAFPGCEYFALFEGQIGVQNDNLPFESPVVEEDGFNLTTSVSM 827	
CC	737 LLLDAALYGLATVLEAVCPQYGIPEPMWNPFRFRSYWCGPRPPKSPAPCTPLD-PKVL 795	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:57:46 ; Search time 27508 Seconds
(without alignments)
11696.929 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctacggggcgccgctccc.....cgtgctggtgaaaaaaaaa 6804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6804	100.0	6804	6	AX644622
2	6762.4	99.4	6791	6	AX686999 Sequence
3	6717.8	98.7	6768	6	AX320362 Sequence
4	6710	98.6	6795	6	AX429481 Sequence
5	6686.4	98.3	6704	9	AF328787
6	6565.6	96.5	6588	9	AF250238 Homo sapi
7	6478.4	95.2	6522	6	AX202218 Sequence
8	6398.4	94.0	6432	6	AX320364 Sequence
9	5928.2	87.1	6027	9	AB055390 Homo sapi
10	5557.2	81.7	6324	6	CQ719367 Sequence
11	5458.2	80.2	5669	6	AX202220 Sequence
12	4381.6	64.4	4653	6	AX644625 Sequence
13	3902.6	57.4	6513	10	AB097814
14	3902.4	57.4	6633	6	AX644619 Sequence
15	3885.8	57.1	6607	6	AX080493 Sequence
16	3885.8	57.1	6607	10	AF287141
17	3303	48.5	5762	6	AX080462 Sequence
18	2861.8	42.1	2911	6	BD234208 ATP-bindi
19	1808.6	26.6	1879	6	AX834707 Sequence

20	1808.6	26.6	1879	9	AK097344	Homo sapi
21	1788.4	26.3	7074	5	AF362377	Gallus ga
22	1740.6	25.6	6786	9	AB055982	Homo sapi
23	1739	25.6	10412	6	CQ785966	Sequence
24	1737.4	25.5	7860	6	AB393480	Sequence
25	1737.4	25.5	7860	6	AX092594	Sequence
26	1737.4	25.5	10442	6	AX060713	Sequence
27	1737.4	25.5	10442	6	AX060892	Sequence
28	1737.4	25.5	10442	9	AF285167	Sequence
29	1735.8	25.5	10474	6	AX060719	Homo sapi
30	1735.8	25.5	10474	6	AX060721	Sequence
31	1735.8	25.5	10474	6	AX060898	Sequence
32	1735.8	25.5	10474	6	AX060900	Sequence
33	1734.2	25.5	6880	6	BD234191	ATP-bindi
34	1734.2	25.5	6880	6	AX253450	Sequence
35	1734.2	25.5	6880	9	HA012376	Homo sapi
36	1734.2	25.5	7260	6	AX253452	Sequence
37	1734.2	25.5	9497	6	AX059973	Sequence
38	1734.2	25.5	9497	9	AF165281	Homo sapi
39	1734.2	25.5	9741	6	AX127830	Sequence
40	1734.2	25.5	9741	6	AX139817	Sequence
41	1734.2	25.5	9741	6	AX351038	Sequence
42	1734.2	25.5	9854	6	AX127831	Sequence
43	1734.2	25.5	9854	6	AX139818	Sequence
44	1725.4	25.4	7878	10	MMABC1	Mus musculu
45	1724.4	25.3	6801	10	AY208182	Rattus no

ALIGNMENTS

RESULT 1
AX644622

LOCUS AX644622 6804 bp DNA linear PAT 27-FEB-2003
DEFINITION Sequence 4 from Patent WO02099108.
ACCESSION AX644622
VERSION AX644622.1 GI:28610637
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Shutter, J. and Ulias, L.
TITLE ATP-binding cassette transporter-like molecules and uses thereof
JOURNAL Patent: WO 02099108-A 4 12-DEC-2002;
Angen, Inc. (US)

FEATURES
source
1..6804
Location/Qualifiers

CDS

/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD68132.1"
/db_xref="GI:28610638"
/translation="MAFWTQMLLLWKNFMYRRRQPVQLLVLLWPLFLFLLVAVRH
SHPLEHEHCHFNKPLPSAGTVPWQLGICNVNNTCPQLTPGEEPRGSRNFNDLSV
SRLLADARTVLGGASARTLAGLKLATLRAARSTAQPTKQSLPEPMLDVAELL
LSRLATSLGLAQAOEPLHLEAEADLAQLLALSLVLRALLQPRGTSGPLE
LLSEALCSVRGSPSTVGPLNWEASDLMELVGOEPESALPDSSLSACSLGALDLS
HPLSLMLRBLKPLILGLKLPADPTPTKLMQAVNETFSELTLRDVREVMGLPR
IFTFMDSNNVAMQLQLLQMDGRRQPRFGRRDHMEALRSFLDPGSGGYSWQAHAD
VHVLGTLGHRVTECLSLDLAAPSEALVSRALQLLAHRFWAGVFLGEGSSDPT
EHTPTDGLGHVRIKIRMDIDVVTNRKIRDFMDPGFAADPLTDLRYVWGFFYLQD
LVRAAVRVLSGANPRAGLYLQOMPYPYVDVDFLRLVLSRLSLTLAWTYSVTLTV
KAVUREKETRLRDMRANGLSRAVLWGLFSLCIGPFLLSAALLVLVLKLDILPYSH
PGVVFILAPAVATVTTQSLLSAFPSRANLAACGLAYFSLYLPVLCVAVNRDLRP
AGGRVATSLSPVAFPGCSLALLEEQGSAQHNWVTRTADVFSLAQVSGULLLD
AALYGLATWYLEAVCPQYIGIPFRRSYPKPPKSPACPTDLPDPLVLEE
APGLSPGVSRSLERKFPQPSQFALRGLSDFYQGHITAFNGHAGKTTLSILSG
LFPFGSGSFIILGHDVRSMAAIRPHLGVCPOYNVLFDMTLTDBHWNVYFGRKLGSAA

VVPEQDRLLQDVGLVSKQSVQTRHLSGCMQKLSVAIAFVGGSQVVLDEPTAGVDP
ASRRGIWELLKRYEGRTLLSLTHLDEAELLGDRVAVAGRLCCCGSPLFLRRHLG
SGYLLTLVAKRLPQTEKADTMEGSDVTQEKNGSQSRVGTQPLLALVQHVHVG
ARLVEPLHELVLPVLTGADHGSFATLRELDLRLAELRLTGTSLSEIEFLKY
VEEADTDMEGSCGQHLCTGLAGLDVTLRLKMPPOSTALENGEPAGSAPETDQSGG
PDVARRVQWALTRQQLALLKRLFLARRSRGLFAQIVLPALFVGLALVFLSLIYPP
FGHPALRLSPTMYGAQVSPFSEDPDGRARLEALLQEALEEPVQVSHRFSFA
PEVPAEAKVLASGNWTPSPSPACQSQPGARRLLPDCPAAAGPPPPCAVTVSGEV
VONLTGRLSDPLVKTVPRLVROGLTKWVNEVRYGGFSLGGRDPLPSGQLGRSV
BELWALLSPPLGGALDRKNTLAWAHSLSLQSLKIWFNNKWHNSVAFVNRASNAI
FTLVLIERVTRKHLQMGGLSPTLYLGNFLWDMCNLYLPACIVLILFAPQORAY
VAPANLPAILLALLLYGMSITPLMYPASFSPFSTAYVLTICNIFLIGINGSMATFV
LELFSDOKLEVSRLKQVELIIPHFCLGRGLDMVNEQAMADAFELGDRQFOSPLR
WEVGNKLLAMVTOGLPFLFTLLQHRSLQLPQVRSLPILGEEDDEVARERERVV
QKATGDLVLRLNRTKYRGORPAVDRLCLIGIPPGECFLGLGVNGHAKTSTFRMVTG
DTLASGEAVLAGHSVAREPSAAHLSNGYCPQSDATPELLTGREHLELLARLGRVPEA
QVAQTAGSLARLVREGRSVMLTSHSMECEALSRALIMVNGRFRCLGSPQHLKGRF
SARRFLWNSLAVVREGSRVPAEAFVAAEPFSGSELREAHGGRLLFQLPGRGRCLARVFGEL
AVHGAHGVDEFSVQTMLEEVFLYFSKQDQKDETHEQEAGVGVDPAPGLQHPRRV
SQFLDDPSTAEVL"

sig_peptide
210...347

ORIGIN

Query Match 100.0%; Score 6804; DB 6; Length 6804;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTCAGGGGCGGCGCGCTCCCTGCTGCTGCTGGCGGAGGGAAGCGCGCAAGAGCTCGG	60
Db	1	CTCAGGGGCGGCGCGCTCCCTGCTGCTGCTGGCGGAGGGAAGCGCGCAAGAGCTCGG	60
Qy	61	AGCCCTTGGAAGAGCTTCCAGGAACCTTCGCTGTGGGATGAAGGAATGAGTTGAGAA	120
Db	61	AGCCCTTGGAAGAGCTTCCAGGAACCTTCGCTGTGGGATGAAGGAATGAGTTGAGAA	120
Qy	121	GGGCGAGGAGTGTCCCGCAGCGCACCGCAGCTTTCAGCCGACCGTTGCTCGACCT	180
Db	121	GGGCGAGGAGTGTCCCGCAGCGCACCGCAGCTTTCAGCCGACCGTTGCTCGACCT	180
Qy	181	CTCTGCTCCGCTCCCGCCAGCTCTACACATGGGCTTCTGGACACAGCTGATGCTGTC	240
Db	181	CTCTGCTCCGCTCCCGCCAGCTCTACACATGGGCTTCTGGACACAGCTGATGCTGTC	240
Qy	241	TCTGGAAGAAATTCATGTATCGCGGAGACAGCGGTCAGCTCTGCTGCTGCTGCTGT	300
Db	241	TCTGGAAGAAATTCATGTATCGCGGAGACAGCGGTCAGCTCTGCTGCTGCTGCTGT	300
Qy	301	GGCTCTCTTCTCTTCT	360
Db	301	GGCTCTCTTCTCTTCT	360
Qy	361	ACCATGAATGCATCTCCCAAAAGCCACATGCGGCGGCGCACCGTCCCTCGGCTCC	420
Db	361	ACCATGAATGCCATCTCCCAAAAGCCACATGCGGCGGCGCACCGTCCCTCGGCTCC	420
Qy	421	AGGCTCTCATCTGTATGTGAACAAACCTGCTTTCGCGAGCTGACACCGGGGAGGAGC	480
Db	421	AGGCTCTCATCTGTATGTGAACAAACCTGCTTTCGCGAGCTGACACCGGGGAGGAGC	480
Qy	481	CCGGCGGCTGAGCAACTTCAAGACTCCCTGCTCTCCGGCTGTAGCCGATGCCCGCA	540
Db	481	CCGGCGGCTGAGCAACTTCAAGACTCCCTGCTCTCCGGCTGTAGCCGATGCCCGCA	540
Qy	541	CTGTGCTGGGAGGGGCGAGTCCCAAGAGCGTGGCTGGCTAGGGAAGCTGATGCCCA	600
Db	541	CTGTGCTGGGAGGGGCGAGTCCCAAGAGCGTGGCTGGCTAGGGAAGCTGATGCCCA	600
Qy	601	CGTGTAGGGCTGACGAGCAAGCGGCGGCTCAACCAACCAAGAGCTTCTCACTGGAAC	660
Db	601	CGTGTAGGGCTGACGAGCAAGCGGCGGCTCAACCAACCAAGAGCTTCTCACTGGAAC	660

Qy	661	CACCCATGCTGGAGTGTGCGGAGCTGTGACGTCACTGTCTGCGCAAGGAATCCCTGGGGT	720
Db	661	CACCCATGCTGGAGTGTGCGGAGCTGTGACGTCACTGTCTGCGCAAGGAATCCCTGGGGT	720
Qy	721	TGSCACTGGGCCAAGCCCAAGAGCCCTTGCACAGCTTGTGGAGGCGCTGAGAGCACTGG	780
Db	721	TGSCACTGGGCCAAGCCCAAGAGCCCTTGCACAGCTTGTGGAGGCGCTGAGAGCACTGG	780
Qy	781	CCAGAGAGCTTCTGGGCGCTGCGGAGCTGTGTGAGCTTGGGCACTGCTCAGAGACCCC	840
Db	781	CCAGAGAGCTTCTGGGCGCTGCGGAGCTGTGTGAGCTTGGGCACTGCTCAGAGACCCC	840
Qy	841	GAGGACACAGCGCCCTTGTGTGACAGGCGCTTGTGAGCTGTGAGGAGCACTA	900
Db	841	GAGGACACAGCGCCCTTGTGTGACAGGCGCTTGTGAGCTGTGAGGAGCACTA	900
Qy	901	GCAGCAAGTGGGCGCTTCCCTCAACTGTGTAACAGGCTAGTCACTGATGAGCTGTGTG	960
Db	901	GCAGCAAGTGGGCGCTTCCCTCAACTGTGTAACAGGCTAGTCACTGATGAGCTGTGTG	960
Qy	961	GGCAGGAGCCAGAAATCGGCGCTGCCAGACAGCACTGAGCCCGGCTGCTCGGAGCTGA	1020
Db	961	GGCAGGAGCCAGAAATCGGCGCTGCCAGACAGCACTGAGCCCGGCTGCTCGGAGCTGA	1020
Qy	1021	TTGGAGCCCTGGACAGCACCCGCTGTCCGCGCTGTCTTGAGACGCGCTGAAAGCTCTGA	1080
Db	1021	TTGGAGCCCTGGACAGCACCCGCTGTCCGCGCTGTCTTGAGACGCGCTGAAAGCTCTGA	1080
Qy	1081	TCCTCGGGAAGCTACTTTTGGACACAGATAACCTTTTACCCGGAAGCTCATGGCCAGG	1140
Db	1081	TCCTCGGGAAGCTACTTTTGGACACAGATAACCTTTTACCCGGAAGCTCATGGCCAGG	1140
Qy	1141	TGAACCGGAGCTTTCGAGGAGCTCACCTGCTGAGGAGTGTCCGAGAGTGTGGGAGATGC	1200
Db	1141	TGAACCGGAGCTTTCGAGGAGCTCACCTGCTGAGGAGTGTCCGAGAGTGTGGGAGATGC	1200
Qy	1201	TGGGACCCCGGATCTTTCACCTTCATGAAAGACAGATTTCCAAATGTGGCATCTCTCAGCGGC	1260
Db	1201	TGGGACCCCGGATCTTTCACCTTCATGAAAGACAGATTTCCAAATGTGGCATCTCTCAGCGGC	1260
Qy	1261	TCCTGAGATGCAAGTGAAGGAAGGACGCCAGACCTGAGGCGCGGACCAATGG	1320
Db	1261	TCCTGAGATGCAAGTGAAGGAAGGACGCCAGACCTGAGGCGCGGACCAATGG	1320
Qy	1321	AGGCTCTGCGATCCTTCTGAGACCTTGGAGCGGCTGCTACAGCTGGCAGGACGACAG	1380
Db	1321	AGGCTCTGCGATCCTTCTGAGACCTTGGAGCGGCTGCTACAGCTGGCAGGACGACAG	1380
Qy	1381	CTGATGTGGGCGACCTTGTGGGCGACCGCTGGGCGAGTGACGAGTGTCTCTTGGACA	1440
Db	1381	CTGATGTGGGCGACCTTGTGGGCGACCGCTGGGCGAGTGACGAGTGTCTCTTGGACA	1440
Qy	1441	AGCTGAGGCGGACCCCTCAGAGGACGCTTGTGTGCGGGCCCTGCAACTGCTCGCGG	1500
Db	1441	AGCTGAGGCGGACCCCTCAGAGGACGCTTGTGTGCGGGCCCTGCAACTGCTCGCGG	1500
Qy	1501	AACATGATTTCTGGGCGGCGTCTTCTTGGGACCTGAGGACTCTTTCAGAGCCCCACAG	1560
Db	1501	AACATGATTTCTGGGCGGCGTCTTCTTGGGACCTGAGGACTCTTTCAGAGCCCCACAG	1560
Qy	1561	AGCACCAACCCAGACCTTGGGCGGCGGCGCAGCTGGGCGATCAAAATCCGATGGAATTG	1620
Db	1561	AGCACCAACCCAGACCTTGGGCGGCGGCGCAGCTGGGCGATCAAAATCCGATGGAATTG	1620
Qy	1621	ACGTGTGTACGAGGACCAATAAGATCAGGGAAGAGTTTGGGACCTTGGCCCGGCGCGG	1680
Db	1621	ACGTGTGTACGAGGACCAATAAGATCAGGGAAGAGTTTGGGACCTTGGCCCGGCGCGG	1680
Qy	1681	ACCCCTGACGACCTGCGCTAGCTGTGGGCGGCTTGTGTACCTGCAAGACTGTGTGG	1740
Db	1681	ACCCCTGACGACCTGCGCTAGCTGTGGGCGGCTTGTGTACCTGCAAGACTGTGTGG	1740
Qy	1741	AGGTCAGCGCTGCGCTGCTCAGCGGCGGCGCAACCCCGGGCGGCGCTCTCTCTGAGC	1800

[illegible]

2821	DB	ACGTCCGCTCCAGACATGGCGCGCCATCCGGCCCCACCTCGGGCGTCTGTCTTCAGTACAAACG	2880
2881	QY	TGCTGTGTTTTCACATGCTGACCGTGGACGACGACGCTCTGTTTCTATGCGGCGGCTGAAGGGTCTC	2940
2881	DB	TGCTGTGTTTTCACATGCTGACCGTGGACGACGACGCTCTGTTTCTATGCGGCGGCTGAAGGGTCTC	2940
2941	QY	TGAGTGCCTGTTAGTGGGCCCCCGAGAGACCGTCTGTCTGCAGGATGTGTGGGCTGTGGTCT	3000
2941	DB	TGAGTGCCTGTTAGTGGGCCCCCGAGAGAGCCGTCTGTCTGCAGGATGTGTGGGCTGTGGTCT	3000
3001	QY	CCAAGCAGAGTGTGCAGACTCGCCACTCTCTCGTGGATGCAACGGAAGCTGTCCGTGG	3060
3001	DB	CCAAGCAGAGTGTGCAGACTCGGCCACTCTCTCGTGGGATGCAACGGAAGCTGTCCGTGG	3060
3061	QY	CCATTGCTTTTGTGGCGGCTCCCAAGTTGTTATCTTGGACGAGCCTACCGCTGGCGTGG	3120
3061	DB	CCATTGCTTTTGTGGCGGCTCCCAAGTTGTTATCTTGGACGAGCCTACCGCTGGCGTGG	3120
3121	QY	ATCTCTGCTTCCCGCCCGGTAATTTGGAGGCTGTCTCTCAAATACCAGAAAGTTCGACGC	3180
3121	DB	ATCTCTGCTTCCCGCCCGGTAATTTGGAGGCTGTCTCTCAAATACCAGAAAGTTCGACGC	3180
3181	QY	TGATCTCTTCCACCCACCTGGATGAGCAGAGCTGTGGGAGACCGTGTGGCTGTGG	3240
3181	DB	TGATCTCTTCCACCCACCTGGATGAGCAGAGCTGTGGGAGACCGTGTGGCTGTGG	3240
3241	QY	TGGCAGGTGGCGCTTGTGCTGTGTGGGCTCCCCACTCTTCTTCGCGCGCTCACCTGGGCT	3300
3241	DB	TGGCAGGTGGCGCTTGTGCTGTGTGGGCTCCCCACTCTTCTTCGCGCGCTCACCTGGGCT	3300
3301	QY	CCGGCTACTACTGACGCTGGTGAAGGCCCGCTGCCCTTGACCAACCAATGAGAAGGCTG	3360
3301	DB	CCGGCTACTACTGACGCTGGTGAAGGCCCGCTGCCCTTGACCAACCAATGAGAAGGCTG	3360
3361	QY	ACACTGACATGAGGGCAGTGTGGACACAGGACGGAAGAAAGAGATGGCAGCCAGGGCA	3420
3361	DB	ACACTGACATGAGGGCAGTGTGGACACAGGACGGAAGAAAGAGATGGCAGCCAGGGCA	3420
3421	QY	GCAGAGTCCGCACTCTCTCAGCTGTCTGGCCCTGTGTACAGCACTGGGTCCCGGGGCACGGC	3480
3421	DB	GCAGAGTCCGCACTCTCTCAGCTGTCTGGCCCTGTGTACAGCACTGGGTCCCGGGGCACGGC	3480
3481	QY	TGGTGGAGAGGCTGCCACACGAGCTGGTGTGTGTCTGCCCTACCGGGTGCCCATGACG	3540
3481	DB	TGGTGGAGAGGCTGCCACACGAGCTGGTGTGTGTCTGCCCTACACGGGTGCCCATGACG	3540
3541	QY	GCAGCTTGCCCACTCTTCCAGAGCTAGACACGGCGCTGGCGGAGCTGAGGCTCACTG	3600
3541	DB	GCAGCTTGCCCACTCTTCCAGAGCTAGACACGGCGCTGGCGGAGCTGAGGCTCACTG	3600
3601	QY	GCTACGGGATCTCCGACACACGCTCCGAGGAGATCTTCTGAAGGTGGTGGAGAGTGTG	3660
3601	DB	GCTACGGGATCTCCGACACACGCTCCGAGGAGATCTTCTGAAGGTGGTGGAGAGTGTG	3660
3661	QY	CTGCGGACACAGATATGGAGGATGGCAGCTGCGGGCAGCACCTATGACAGGCAATTGCTG	3720
3661	DB	CTGCGGACACAGATATGGAGGATGGCAGCTGCGGGCAGCACCTATGACAGGCAATTGCTG	3720
3721	QY	GCCTAGAGTAACCTTCGGGCTCAAGATGCGGCCACAGGAGACAGCGCTGAGGAACGGGG	3780
3721	DB	GCCTAGAGTAACCTTCGGGCTCAAGATGCGGCCACAGGAGACAGCGCTGAGGAACGGGG	3780
3781	QY	AACCAAGTGGGTTCAGCCCCAGAGACTGACACGGGCTCTGGGCCAGACGCGCTGGGCGGG	3840
3781	DB	AACCAAGTGGGTTCAGCCCCAGAGACTGACACGGGCTCTGGGCCAGACGCGCTGGGCGGG	3840
3841	QY	TACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCCCTGCTTCTCAAGCGCTTTCCTGC	3900
3841	DB	TACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCCCTGCTTCTCAAGCGCTTTCCTGC	3900
3901	QY	TTGCCCGCGCAGCGCGCGGCTGTTTGGCCAGATCGTGTGCTGTGCTGCTTTTGTGG	3960
3901	DB	TTGCCCGCGCAGCGCGCGGCTGTTTGGCCAGATCGTGTGCTGTGCTGCTTTTGTGG	3960

Qy	3961	GCCTGGCCCTCGTGTTCAGCCTCATCGTGCTCTCTTTTCGGGCACTACCCCGCTCTGCGC	4020
Db	3961	GCCTGGCCCTCGTGTTCAGCCTCATCGTGCTCTCTTTTCGGGCACTACCCCGCTCTGCGC	4020
Qy	4021	TCAGTCCCAACATGTACGGTGCTCAGGTGTCTTCTTCAGTGAGGACGCCCCAGGGGACC	4080
Db	4021	TCAGTCCCAACATGTACGGTGCTCAGGTGTCTTCTTCTTCAGTGAGGACGCCCCAGGGGACC	4080
Qy	4081	CTGACGTGCCCGCTGCTCAGAGCGCTGCTGCAGGAGGAGGACTGGAGAGCCCCCAG	4140
Db	4081	CTGAGCTGCCCGCTGCTCAGAGCGCTGCTGCAGGAGGAGGACTGGAGAGCCCCCAG	4140
Qy	4141	TGCAGCATAGCTCCCA CAGGTTCTGGCA C CAGAAAGTTCCTGTGTAAGTGGCCAAAGTCT	4200
Db	4141	TGCAGCATAGCTCCCA CAGGTTCTCGGCA C CAGAAAGTTCCTGTGTAAGTGGCCAAAGTCT	4200
Qy	4201	TGGCCAGTGGCAACTGAGACCCAGAGTCTCCATCCCGCAGCTGCCAGTGTAGCAGGCCG	4260
Db	4201	TGGCCAGTGGCAACTGAGACCCAGAGTCTCCATCCCGCAGCTGCCAGTGTAGCAGGCCG	4260
Qy	4261	GTGCCCGCGCTGCTGCCGACTGCCCGCTGCAGCTGTGTGTGTCCCGCTCCGCCCCCAGG	4320
Db	4261	GTGCCCGCGCTGCTGCCGACTGCCCGCTGCAGCTGTGTGTGTCCCGCTCCGCCCCCAGG	4320
Qy	4321	CAGTGA CCGGCTCTGGGGAAGTGTT CAGAA C CTGA CAGGCCGGAA CCTGTCTGACTTCC	4380
Db	4321	CAGTGA CCGGCTCTGGGGAAGTGTT CAGAA C CTGA CAGGCCGGAA CCTGTCTGACTTCC	4380
Qy	4381	TGTTCAAGACCTACCCGCTGTGTGGCGCAGGGCTGAAGACTTAAGAAGTGGGTGAATG	4440
Db	4381	TGTTCAAGACCTACCCGCGCTGTGTGGCGCAGGGCTGAAGACTTAAGAAGTGGGTGAATG	4440
Qy	4441	AGGT CAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCTGCCCTCGGGCCAAAG	4500
Db	4441	AGGT CAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCTGCCCTCGGGCCAAAG	4500
Qy	4501	AGTTGGGGCGCTCAGTGGAGAGTTGTGGGGCTGTGAGTCCCGCTGCTGGCGGGGCC	4560
Db	4501	AGTTGGGGCGCTCAGTGGAGAGTTGTGGGGCTGTGAGTCCCGCTGCTGGCGGGGCC	4560
Qy	4561	TCGACCGTGTCTCAAAAACCTCAGCCTGGGGCTCAGAGCTTGATGTCTCAGGACAGTC	4620
Db	4561	TCGACCGTGTCTCAAAAACCTCAGCCTGGGGCTCAGAGCTTGATGTCTCAGGACAGTC	4620
Qy	4621	TCAAGATCTGGTTCAACAACAAAGGCTGGCATCCATGGTGGCTTTGTCAACCGAGCCA	4680
Db	4621	TCAAGATCTGGTTCAACAACAAAGGCTGGCATCCATGGTGGCTTTGTCAACCGAGCCA	4680
Qy	4681	GCAACGCATCTCTCGTGCTCACCTGCCCCCAGGCCCGGCCAGCCACGCCACAGCATCA	4740
Db	4681	GCAACGCATCTCTCGTGCTCACCTGCCCCCAGGCCCGGCCAGCCACGCCACAGCATCA	4740
Qy	4741	CCACACTCAACACCCCTTGAACTCAACAGGAGCAGCTGTCTGAGGCTGCATGTATGG	4800
Db	4741	CCACACTCAACACCCCTTGAACTCAACAGGAGCAGCTGTCTGAGGCTGCATGTATGG	4800
Qy	4801	CCTCTCGGTGGACGTCTCGTCTCCATCTGTGTGTGCTTTTGCCATGTCTTTGTCCCGG	4860
Db	4801	CCTCTCGGTGGACGTCTCGTCTCCATCTGTGTGTGCTTTTGCCATGTCTTTGTCCCGG	4860
Qy	4861	CCAGTTCACCTTGTCTCATTTAGAGAGGAGTCA CCGAGCCCAAGCACTCGACGTCA	4920
Db	4861	CCAGTTCACCTTGTCTCATTTAGAGAGGAGTCA CCGAGCCCAAGCACTCGACGTCA	4920
Qy	4921	TGGGGGGCTCTGCCCCACCCCTCTACTGGCTTGGCAACTTTTCTCTGGGACATGTGTAAC	4980
Db	4921	TGGGGGGCTCTCTGCCCCACCCCTCTACTGGCTTGGCAACTTTCTCTGGGACATGTGTAAC	4980
Qy	4981	ACTTGTGCGCAGCATGATCGTGTGTCTCATCTTTCTTGGCTTCCAGCAGAGGCATATG	5040
Db	4981	ACTTGTGCGCAGCATGATCGTGTGTCTCATCTTTCTTGGCTTCCAGCAGAGGCATATG	5040

QY	5041	TGGCCCTGCGAACCTGCTGCTCTCTCTGCTGTGTGCTACTACTGTATGGCTGTGCTGATCA	5100
DB	5041	TGGCCCTGCGAACCTGCTGCTCTCTCTGCTGTGTGCTACTACTGTATGGCTGTGCTGATCA	5100
QY	5101	CACGCTCATGTATCCCAAGCTCCTTCTTCTTCTCCGTGCCACGACACAGCTATGTGTGTC	5160
DB	5101	CACGCTCATGTATCCCAAGCTCCTTCTTCTTCTTCTCCGTGCCACGACACAGCTATGTGTGTC	5160
QY	5161	TCACCTGCATAAACCCTCTTTATTTGGCATCAATGGAAAGCATGGCCACCTTTGTGCTTGAGC	5220
DB	5161	TCACCTGCATAAACCCTCTTTATTTGGCATCAATGGAAAGCATGGCCACCTTTGTGCTTGAGC	5220
QY	5221	TCCTTCTGTGATCAGAAAGCTGCAAGAGGTGAGCCGGATCTTTGAAACAGGTCTTCTCTATCT	5280
DB	5221	TCCTTCTGTGATCAGAAAGCTGCAAGAGGTGAGCCGGATCTTTGAAACAGGTCTTCTCTATCT	5280
QY	5281	TCCCCCACTTCTGCTTTGGGCGGGGGCTCAATTGACATGTTGCCGNAACCAAGCCATGGCTG	5340
DB	5281	TCCCCCACTTCTGCTTTGGGCGGGGGCTCAATTGACATGTTGCCGNAACCAAGCCATGGCTG	5340
QY	5341	ATGCTTTTGTAGCGCTTGGGAGACAGGCAGTTTCAGTCAACCCTCGCTCGGTGGGAGTGGTCG	5400
DB	5341	ATGCTTTTGTAGCGCTTGGGAGACAGGCAGTTTCAGTCAACCCTCGCTCGGTGGGAGTGGTCG	5400
QY	5401	GCAAGAACCTCTTTGGCCATGGTGATACAGGGGCCCTCTTCTTCTTTCACACTACTGC	5460
DB	5401	GCAAGAACCTCTTTGGCCATGGTGATACAGGGGCCCTCTTCTTCTTTCACACTACTGC	5460
QY	5461	TGACAGCACGAAGCCAACTCTTGCCACAGCCGAGGTGAGGTCTGTGCCACTCTCTGGGAG	5520
DB	5461	TGACAGCACGAAGCCAACTCTTGCCACAGCCGAGGTGAGGTCTGTGCCACTCTCTGGGAG	5520
QY	5521	AGGAGGACGAGGATGTAGCCCGCTGAACCGGAGCGGGTGTCAAAGGAGCCACCACGAGGGG	5580
DB	5521	AGGAGGACGAGGATGTAGCCCGCTGAACCGGAGCGGGTGTCAAAGGAGCCACCACGAGGGG	5580
QY	5581	ATGTGTTGGTCTGAGGAACTTGACAAAGTATACCGTGGGCAGAGATGCCAGCTGTTG	5640
DB	5581	ATGTGTTGGTCTGAGGAACTTGACAAAGTATACCGTGGGCAGAGATGCCAGCTGTTG	5640
QY	5641	ACCGCTTGTGCTGGGGATTTCCCTCTGGTGAGTGTTTTGGGCTGCTGGGTGTGAATGGAG	5700
DB	5641	ACCGCTTGTGCTGGGGATTTCCCTCTGGTGAGTGTTTTGGGCTGCTGGGTGTGAATGGAG	5700
QY	5701	CAGGGAAGACGTCCACGCTTTTCGATGTGACGGGGACATTTGGCCAGCAGCGGGCGAGG	5760
DB	5701	CAGGGAAGACGTCCACGCTTTTCGATGTGACGGGGACATTTGGCCAGCAGCGGGCGAGG	5760
QY	5761	CTGTGCTGGCAGGCCACAGCGTGGCCGGGAAACCAAGTGTGCGCACTTCAGCATGGAT	5820
DB	5761	CTGTGCTGGCAGGCCACAGCGTGGCCGGGAAACCAAGTGTGCGCACTTCAGCATGGAT	5820
QY	5821	ACTGCCCTCAATCGATGCCATCTTTGAGCTGTGACGGGCCGGGAGCACTTGAGAGCTGC	5880
DB	5821	ACTGCCCTCAATCGATGCCATCTTTGAGCTGTGACGGGCCGGGAGCACTTGAGAGCTGC	5880
QY	5881	TTGCGCGCTGCGGGGTGCCGAGGCCAGGTTGCCAGACCGCTGGCTCAGGCCCTGG	5940
DB	5881	TTGCGCGCTGCGGGGTGCCGAGGCCAGGTTGCCAGACCGCTGGCTCAGGCCCTGG	5940
QY	5941	CGCGCTGGGACTCTCATGTGTAAGCAGACCGGCTTCAGGCACCTACAGCGGAGGAACA	6000
DB	5941	CGCGCTGGGACTCTCATGTGTAAGCAGACCGGCTTCAGGCACCTACAGCGGAGGAACA	6000
QY	6001	AACGCAAGCTGGGACGGCCCTGTGTTGGGAGCCAGCGCTGTGTTCTTGACG	6060
DB	6001	AACGCAAGCTGGGACGGCCCTGTGTTGGGAGCCAGCGCTGTGTTCTTGACG	6060
QY	6061	AGCCGACACAGGCATGGAACCCAGCGCGCGGCTTCTTCTTGGGAACAGCTTTTGCCCG	6120
DB	6061	AGCCGACACAGGCATGGAACCCAGCGCGCGGCTTCTTCTTGGGAACAGCTTTTGCCCG	6120
QY	6121	TGTTGGGAGGGCCGTTTCAGTGATGCTCACTCCCATAGCATGGAGAGTGTGAAGCGC	6180

972 GAATCCGCCCTGCAGACAGCAGCTGAGCCCGCCCTGCTCGGAGCTGATTGGAGCCCTG 1031
1032 GACAGCCACCCGCTGTCCCGCTGCTCTGAGAGCGCTGAAGCCTCTGATCTCTCGGGAAG 1091
1020 GACAGCCACCCGCTGTCCCGCTGCTCTGAGAGCGCTGAAGCCTCTGATCTCTCGGGAAG 1079
1092 CTAATCTTTGACACAGATACACCTTTTACCGGAAGCTATGCGGCAGAGTGAACCGGACC 1151
1080 CTAATCTTTGACACAGATACACCTTTTACCGGAAGCTATGCGGCAGAGTGAACCGGACC 1139
1152 TTCGAGGAGCTACCTCTGCTGAGGAGTATGCGGAGGTGCGGAGATGCTGGGACCCCGG 1211
1140 TTCGAGGAGCTACCTCTGCTGAGGAGTATGCGGAGGTGCGGAGATGCTGGGACCCCGG 1199
1212 ATCTTCACTTATGAACGACAGTTCCTCAATGTGGCCATGCTGAGCGGCTCTCTGCAGATG 1271
1200 ATCTTCACTTATGAACGACAGTTCCTCAATGTGGCCATGCTGAGCGGCTCTCTGCAGATG 1259
1272 CAGGATGAAGGAAGAGCGGACCCAGACCTGGAGGCGGGACACATGAGGCGCTCTGCGA 1331
1260 CAGGATGAAGGAAGAGCGGACCCAGACCTGGAGGCGGGACACATGAGGCGCTCTGCGA 1319
1332 TCCTTTCTGGACCTTGGAGCGGTGGCTACAGCTGCGAGGAGCACAACGCTGATGTGGG 1391
1320 TCCTTTCTGGACCTTGGAGCGGTGGCTACAGCTGCGAGGAGCACAACGCTGATGTGGG 1379
1392 CACTTGTGGGACCGCTGGGCGGAGTGAAGGAGTGCCTGTCTTTGGACAAGCTGGAGGCG 1451
1380 CACTTGTGGGACCGCTGGGCGGAGTGAAGGAGTGCCTGTCTTTGGACAAGCTGGAGGCG 1439
1452 GCACCTTCAGAGCAGCCCTGTGTGCGGGGCCCTGCAACTGCTGCGGGAACATCGATTTC 1511
1440 GCACCTTCAGAGCAGCCCTGTGTGCGGGGCCCTGCAACTGCTGCGGGAACATCGATTTC 1499
1512 TGGGCGGGCTGTCTTTCTGGGACCTGAGGACTCTTTACAGCCCAAGAGCAGCAGCAGC 1571
1500 TGGGCGGGCTGTCTTTCTGGGACCTGAGGACTCTTTACAGCCCAAGAGCAGCAGCAGC 1559
1572 CCAGACCTTGGGCGGCGGACCGTGGCGCATCAAAATCCGATGGACATTTGACGTGGTCAAG 1631
1560 CCAGACCTTGGGCGGCGGACCGTGGCGCATCAAAATCCGATGGACATTTGACGTGGTCAAG 1619
1632 AGSACCAATGAATCAGGAGCAGGTTTGGGACCTTGGGACCTTGGGACCGGCGGACCCCTGACC 1691
1620 AGSACCAATGAATCAGGAGCAGGTTTGGGACCTTGGGACCGGCGGACCGGCGGACCCCTGACC 1679
1692 GACCTGCGCTACGCTGTGGGCGGCTTCTGTGTACTGCAAGACCTGGTGGAGCGTGCAGCC 1751
1680 GACCTGCGCTACGCTGTGGGCGGCTTCTGTGTACTGCAAGACCTGGTGGAGCGTGCAGCC 1739
1752 GTCGCGGTGCTCAGCGGCGGCGCAACCCCGGGCGGCGCTCTTACCTGAGCAGAGTGCCTAT 1811
1740 GTCGCGGTGCTCAGCGGCGGCGCAACCCCGGGCGGCGCTCTTACCTGAGCAGAGTGCCTAT 1799
1812 CCGTGTATGTGAAGCAGCTGTTCCTGCTGTGCTGAGCGGCTGAGCGGCTGCTGCTCTTCCTG 1871
1800 CCGTGTATGTGAAGCAGCTGTTCCTGCTGTGCTGAGCGGCTGAGCGGCTGCTGCTCTTCCTG 1859
1872 ACCTGGCGCTTGAATCTTACTCCGTGACACTGACAGTGAAGGCGCTGTGGGGAAGAGGAG 1931
1860 ACCTGGCGCTTGAATCTTACTCCGTGACACTGACAGTGAAGGCGCTGTGGGGAAGAGGAG 1919
1932 ACAGCGCTGCGGACACCAATGCGCGCATGAGGCTCAGCGCGGCTGCTGCTGAGGCTAGGC 1991
1920 ACAGCGCTGCGGACACCAATGCGCGCATGAGGCTCAGCGCGGCTGCTGCTGAGGCTAGGC 1979
1992 TGGTTCTCAGCTGCTCGGGCCCTTCTGCTCAGCGCGGCTGCTGCTGAGGCTGCTC 2051
1980 TGGTTCTCAGCTGCTCGGGCCCTTCTGCTCAGCGCGGCTGCTGCTGAGGCTGCTC 2039
2052 AAGCTGGGGACATCTCCCTACAGCCACCCGGGCGTGGTCTTCTCTGTTCTTGGCAGCC 2111

2040 AAGCTGGGGACATCTCCCTACAGCCACCCGGGCGTGGTCTTCTCTGTTCTTGGCAGCC 2099
2112 TTCGCGGTGGCCACCGGTGACCCAGAGCTTCTGCTCAGCGCCTTCTTCTCCCGCGCAAC 2171
2100 TTCGCGGTGGCCACCGGTGACCCAGAGCTTCTGCTCAGCGCCTTCTTCTCCCGCGCAAC 2159
2172 CTGGCTTGGCGCTGCGGCGGCTTGGCTACTTCTCCCTCTACTCTGCGCTACGCTGCTGT 2231
2160 CTGGCTTGGCGCTGCGGCGGCTTGGCTACTTCTCCCTCTACTCTGCGCTACGCTGCTGT 2219
2232 GTGGCTTGGCGGACCCGCTGCGCGGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 2291
2220 GTGGCTTGGCGGACCCGCTGCGCGGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 2279
2292 GTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2351
2280 GTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2339
2352 CAGTGGCAAAAGTGGGACCCGCTACGCGGCTACGCGAGAGCTCTTACGCTGCGGCTGCTCT 2411
2340 CAGTGGCAAAAGTGGGACCCGCTACGCGGCTACGCGAGAGCTCTTACGCTGCGGCTGCTCT 2399
2412 GGCCTTCTGCTGTGACCGCGCTCTACGCGCTCGCCACCTGTGTACCTGGAAGCTGTG 2471
2400 GGCCTTCTGCTGTGACCGCGCTCTACGCGCTCGCCACCTGTGTACCTGGAAGCTGTG 2459
2472 TGCACAGGCGAGTGAAGGATTCCTGAAACCATGGAATTTCTTTTCGAGGAGCTACTGG 2531
2460 TGCACAGGCGAGTGAAGGATTCCTGAAACCATGGAATTTCTTTTCGAGGAGCTACTGG 2519
2532 TCGGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2591
2520 TCGGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2579
2592 CTGGTAGAAGAGGCGACCGCGCGCTGAGTCTCTGGGCTATCCGTTTCGAGCCTGGAGAAG 2651
2580 CTGGTAGAAGAGGCGACCGCGCGCTGAGTCTCTGGGCTCTCGCTTCGAGCCTGGAGAAG 2639
2652 CGCTTCTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2711
2640 CGCTTCTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2699
2712 CACATCACCGCTTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2771
2700 CACATCACCGCTTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2759
2772 AGTGGCGCTTCTCCACCGAGTGGTGGCTCTGCTTCTATCTGGGCGGCGGCGGCGGCGGCGG 2831
2760 AGTGGCGCTTCTCCACCGAGTGGTGGCTCTGCTTCTATCTGGGCGGCGGCGGCGGCGGCGG 2819
2832 AGATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2891
2820 AGCATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2879
2892 ATGCTGACCGTGGAGCGAGCAGCTGCTGTTCTATGGCGGCTGAAGGCTCTGAGTGGCGCT 2951
2880 ATGCTGACCGTGGAGCGAGCAGCTGCTGTTCTATGGCGGCTGAAGGCTCTGAGTGGCGCT 2939
2952 GTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3011
2940 GTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2999
3012 GTGAGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3071
3000 GTGAGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3059
3072 GTGGGCGGCTCCAGTGTGTATCTTGGAGCGGCTTACGCGGCTGGCGTGTGCTGCTTCC 3131
3060 GTGGGCGGCTCCAGTGTGTATCTTGGAGCGGCTTACGCGGCTGGCGTGTGCTGCTTCC 3119
3132 CGCGCGGCTATTTGGGAGCTGCTGCTCAAAATACCGAGAGGCTCGCACGCTGATCTCTCC 3191

Db 3120 CCGCGGGTATTGGAGTGTGCTCAATAACCGAAGGTGCGACGCTGATCCTCC 3179
Qy 3192 ACCCAACCTGGATGAGCAGAGCTGCTGGAGACCGTGTGGCTGTGTGGTGGCAGGTGGC 3251
Db 3180 ACCCAACCTGGATGAGCAGAGCTGCTGGAGACCGTGTGGCTGTGTGGTGGCAGGTGGC 3239
Qy 3252 CGCTTGTGCTGTGGTGTCCCACTCTTCCTGCGCGCTCAGCTGGGCTCGGGCTACTAC 3311
Db 3240 CGCTTGTGCTGTGGATCCCACTCTTCCTGCGCGCTCAGCTGGGCTCGGGCTACTAC 3299
Qy 3312 CTGACGCTGGTGAAGCGCGCTGCCCTGACCAACCAATGAAAGGCTGACACTGACATG 3371
Db 3300 CTGACGCTGGTGAAGCGCGCTGCCCTGACCAACCAATGAAAGGCTGACACTGACATG 3359
Qy 3372 GAGGCACTGTGGACACCAAGGAGAAAGAAATGGCAGCAGCGGAGCAGAGTCCGC 3431
Db 3360 GAGGCACTGTGGACACCAAGGAGAAAGAAATGGCAGCAGCGGAGCAGAGTCCGC 3419
Qy 3432 ACTCTCAGCTGCTGGCCCTGGTACAGCACTGGGTGCCGGGCA CGGCTGTGGAGAG 3491
Db 3420 ACTCTCAGCTGCTGGCCCTGGTACAGCACTGGGTGCCGGGCA CGGCTGTGGAGAG 3479
Qy 3492 CTGCCACACAGCTGCTGCTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCC 3551
Db 3480 CTGCCACACAGCTGCTGCTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCC 3539
Qy 3552 ACACCTTCCGAGACTAGACACGCGGCTGGCGAGCTCAGGCTCACTGGGCTACGGGATC 3611
Db 3540 ACACCTTCCGAGACTAGACACGCGGCTGGCGAGCTCAGGCTCACTGGGCTACGGGATC 3599
Qy 3612 TCCGACACAGCTCGAGAGATCTTCCTGAAGGTGGTGGAGAGTGTGCTGCGGACACA 3671
Db 3600 TCCGACACAGCTCGAGAGATCTTCCTGAAGGTGGTGGAGAGTGTGCTGCGGACACA 3659
Qy 3672 GATATGGAGATGGCAGCTGGGCGAGCACCTATGACAGGCATGCTGGGCTAGACGTA 3731
Db 3660 GATATGGAGATGGCAGCTGGGCGAGCACCTATGACAGGCATGCTGGGCTAGACGTA 3719
Qy 3732 ACCCTGCGGCTCAAGATCCGCCACAGGAGACAGCGCTGGAGAACCGGGAACCAAGCTGGG 3791
Db 3720 ACCCTGCGGCTCAAGATCCGCCACAGGAGACAGCGCTGGAGAACCGGGAACCAAGCTGGG 3779
Qy 3792 TCAGCCCCAGAGACTGACAGGGCTCTGGGCGAGACCGCTGGGCGGGTACAGGGCTGG 3851
Db 3780 TCAGCCCCAGAGACTGACAGGGCTCTGGGCGAGACCGCTGGGCGGGTACAGGGCTGG 3839
Qy 3852 GCACGTACCCGCGAGCAGCTCCAGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCGCGC 3911
Db 3840 GCACGTACCCGCGAGCAGCTCCAGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCGCGC 3899
Qy 3912 AGCGCCGCGGCTGTTCCGCCAGATCGTGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3971
Db 3900 AGCGCCGCGGCTGTTCCGCCAGATCGTGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3959
Qy 3972 GTGTTACGCTCATGTCCTCTTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCCAACC 4031
Db 3960 GTGTTACGCTCATGTCCTCTTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCCAACC 4019
Qy 4032 ATGTACGCTGCTCAGGTGCTCTTCTTCAGTAGAGACGCCACAGGGACCTTGAGGTGCC 4091
Db 4020 ATGTACGCTGCTCAGGTGCTCTTCTTCAGTAGAGACGCCACAGGGACCTTGAGGTGCC 4079
Qy 4092 CGGCTGCTCGAGGCGCTCTCGAGGAGCAGGACTGGAGGAGCCGCCAGTGCAGCATAGC 4151
Db 4080 CGGCTGCTCGAGGCGCTCTCGAGGAGCAGGACTGGAGGAGCCGCCAGTGCAGCATAGC 4139
Qy 4152 TCCCAAGGTTCTCGGCAACCAAGAGTTCCTGCTGAAGTGGCAAGGCTTGTGGCCAGTGGC 4211
Db 4140 TCCCAAGGTTCTCGGCAACCAAGAGTTCCTGCTGAAGTGGCAAGGCTTGTGGCCAGTGGC 4199
Qy 4212 AACTGGACCCAGAGTCTCCATCCCGACGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGC 4271
Db 4200 AACTGGACCCAGAGTCTCCATCCCGACGCTGCCAGTGTAGCCCGGCGCGGTGCCCGGCGC 4259

Qy 4272 CTGCTGCCGAGCTGCCCGGCTGCAGCTGGTGGTCCCTCCGCCCCAGCAGTGCAGCGC 4331
Db 4260 CTGCTGCCGAGCTGCCCGGCTGCAGCTGGTGGTCCCTCCGCCCCAGCAGTGCAGCGC 4319
Qy 4332 TCTGGGGAAGTGGTTAGAACTGTGACAGCGCGGAACCTGTGTGACTTCTGTGTCAAGACC 4391
Db 4320 TCTGGGGAAGTGGTTAGAACTGTGACAGCGCGGAACCTGTGTGACTTCTGTGTCAAGACC 4379
Qy 4392 TACC CGGCTTGTGCGCAGCGCTGAAGACTAAGAAAGTGGTGAATGAGTCAAGTAC 4451
Db 4380 TACC CGGCTTGTGCGCAGCGCTGAAGACTAAGAAAGTGGTGAATGAGTCAAGTAC 4439
Qy 4452 GGAGGCTTCTCGCTGGGGGCGCGAGACCCAGGCTTCCCTCGGGCCAAAGAGTTGGGCGCGC 4511
Db 4440 GGAGGCTTCTCGCTGGGGGCGCGAGACCCAGGCTTCCCTCGGGCCAAAGAGTTGGGCGCGC 4499
Qy 4512 TCAGTGGAGGAGTGTGGGCGCTGTGAGTCCCTGCTGGGCGGCGCTTCGACCGTGTGTC 4571
Db 4500 TCAGTGGAGGAGTGTGGGCGCTGTGAGTCCCTGCTGGGCGGCGCTTCGACCGTGTGTC 4559
Qy 4572 CTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGG 4631
Db 4560 CTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGG 4619
Qy 4632 TTCAACAAACAAAGGCTGGCACTCCATGTGGGCTTTTGTCAACCGAGCCAGCAACGCAATC 4691
Db 4620 TTCAACAAACAAAGGCTGGCACTCCATGTGGGCTTTTGTCAACCGAGCCAGCAACGCAATC 4679
Qy 4692 CTCCGTGCTCACTTGGCCCCAGGGCGCGCCAGCCGCGCCACAGCATCACACACTCAAC 4751
Db 4680 CTCCGTGCTCACTTGGCCCCAGGGCGCGCCAGGGCGCGCCAGCATCACACACTCAAC 4739
Qy 4752 CACCCCTGAACTCACCAAGGAGCAGCTGCTGAGGCTGACATGAGCTTCCCTCGGTCG 4811
Db 4740 CACCCCTGAACTCACCAAGGAGCAGCTGCTGAGGCTGACATGAGCTTCCCTCGGTCG 4799
Qy 4812 GACGCTCTCGTCTCCATCTGTGTGCTTTTGGCCATGTCTTTGTCGCGGCGAGCTTCACT 4871
Db 4800 GACGCTCTCGTCTCCATCTGTGTGCTTTTGGCCATGTCTTTGTCGCGGCGAGCTTCACT 4859
Qy 4872 CTGTGCTCATATGAGGAGCAGTCA CCGAGCCCAAGCACTGCGAGCTCATGGGGGCGCTG 4931
Db 4860 CTGTGCTCATATGAGGAGCAGTCA CCGAGCCCAAGCACTGCGAGCTCATGGGGGCGCTG 4919
Qy 4932 TCCCCCACCCCTACTGCTTGGCAA CTTTCTCTGSGGACATGTGTAACTACTTGTGTGCA 4991
Db 4920 TCCCCCACCCCTACTGCTTGGCAA CTTTCTCTGSGGACATGTGTAACTACTTGTGTGCA 4979
Qy 4992 GCATGCATCGTGTGCTCATCTTTTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGCC 5051
Db 4980 GCATGCATCGTGTGCTCATCTTTTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGCC 5039
Qy 5052 AACCTGCTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATCACAACCGCTCATG 5111
Db 5040 AACCTGCTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATCACAACCGCTCATG 5099
Qy 5112 TACCAGGCTCTTCTTCTCTCGTCCGCGCAGCAGCTATGTGGTGTGCTCACCTGCATA 5171
Db 5100 TACCAGGCTCTTCTTCTCTCGTCCGCGCAGCAGCTATGTGGTGTGCTCACCTGCATA 5159
Qy 5172 AACCTCTTTATTTGGCATCAATGGAAGCATGGCCACCTTTTGTGCTTGAAGCTCTTCTCTGAT 5231
Db 5160 AACCTCTTTATTTGGCATCAATGGAAGCATGGCCACCTTTTGTGCTTGAAGCTCTTCTCTGAT 5219
Qy 5232 CAGAAGCTGACGAGGTGAGCCGGATCTTGAAAACAGGCTTTCCTTATCTTCCCCACCTTC 5291
Db 5220 CAGAAGCTGACGAGGTGAGCCGGATCTTGAAAACAGGCTTTCCTTATCTTCCCCACCTTC 5279
Qy 5292 TGCTTGGGCGGGGCTCATTTGACATGGTGGGAAACCAAGGCATGGCTGATGCCCTTTCAG 5351
Db 5280 TGCTTGGGCGGGGCTCATTTGACATGGTGGGAAACCAAGGCATGGCTGATGCCCTTTCAG 5339

QY 5352 CGCTTGGGAGACAGGACGAGTTCCAGTCAACCTCGCTGGGAGGTGGTGGCAAGAACCTC 5411
DB 5340 CGCTTGGGAGACAGGACGAGTTCCAGTCAACCTCGCTGGGAGGTGGTGGCAAGAACCTC 5399
QY 5412 TTGGCCATGGTGATACAGGGGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5471
DB 5400 TTGGCCATGGTGATACAGGGGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5459
QY 5472 AGCCAACTCTGTCACAGCCAGCGGTGAGTCTCTGCACTCTCTGGGAGAGGAGGAG 5531
DB 5460 AGCCAACTCTGTCACAGCCAGCGGTGAGTCTCTGCACTCTCTGGGAGAGGAGGAG 5519
QY 5532 GATGTAGCCGCTGAACCGGAGCGGTGTCACAGGAGCACACCCAGAGGGGAGTGTGGTG 5591
DB 5520 GATGTAGCCGCTGAACCGGAGCGGTGTCACAGGAGCACACCCAGAGGGGAGTGTGGTG 5579
QY 5592 CTGAGGAATCTGACCAAGGTATACCGTGGGAGAGGATGCCAGTGTGTGACCGCTTGTGC 5651
DB 5580 CTGAGGAATCTGACCAAGGTATACCGTGGGAGAGGATGCCAGTGTGTGACCGCTTGTGC 5639
QY 5652 CTGGGGATTCCTCTGTGAGTGTGGGCTGTGGTGTGTAATGGAGAGGAGGAGAG 5711
DB 5640 CTGGGGATTCCTCTGTGAGTGTGGGCTGTGGTGTGTAATGGAGAGGAGGAGAG 5699
QY 5712 TCACAGTTTCGATGTGACGGGGACACATTCGCCAGAGGCGGAGGCTGTGCTGGCA 5771
DB 5700 TCACAGTTTCGATGTGACGGGGACACATTCGCCAGAGGCGGAGGCTGTGCTGGCA 5759
QY 5772 GGCCACAGCGTGGCCCGGGAACCCAGTGTGCGCACCTCAGCATGGGATGTCGCCCTCAA 5831
DB 5760 GGCCACAGCGTGGCCCGGGAACCCAGTGTGCGCACCTCAGCATGGGATGTCGCCCTCAA 5819
QY 5832 TCCGATGCCATCTTTGAGCTGTGAACGGGCGGAGCACCTGAGAGTGTGTGGCGGCTG 5891
DB 5820 TCCGATGCCATCTTTGAGCTGTGAACGGGCGGAGCACCTGAGAGTGTGTGGCGGCTG 5879
QY 5892 CGGGTGTCCGGAGGCGGAGTGGCCAGACCGCTGGCTCAGGCGCTGGCGCTGGGA 5951
DB 5880 CGGGTGTCCGGAGGCGGAGTGGCCAGACCGCTGGCTCAGGCGCTGGCGCTGGGA 5939
QY 5952 CTCTCATGATGACGAGACCGGCTTGACGAGCACCTTACAGCGAGGAGAACAAACCAAGCTG 6011
DB 5940 CTCTCATGATGACGAGACCGGCTTGACGAGCACCTTACAGCGAGGAGAACAAACCAAGCTG 5999
QY 6012 GCGACGCGCTTGGCGTGTGGGACCCAGCGTGTGTTCTGAGAGGAGCCAGCAC 6071
DB 6000 GCGACGCGCTTGGCGTGTGGGACCCAGCGTGTGTTCTGAGAGGAGCCAGCAC 6059
QY 6072 GGCATGACCCAGCGCGCGCGCTTCTTTGGAAACAGCGCTTTTGGCGGTGTGGGAG 6131
DB 6060 GGCATGACCCAGCGCGCGCGCTTCTTTGGAAACAGCGCTTTTGGCGGTGTGGGAG 6119
QY 6132 GGCCTGTCAGTGATGCTCACTCCCATAGCATGGAGAGTGTGAAGCGCTCTGCTCGCGC 6191
DB 6120 GGCCTGTCAGTGATGCTCACTCCCATAGCATGGAGAGTGTGAAGCGCTCTGCTCGCGC 6179
QY 6192 CTAGCCATCATGTGTAAATGGCGGTTCCGCTGCTGGGAGCCCGCAACATCTCAAGGCG 6251
DB 6180 CTAGCCATCATGTGTAAATGGCGGTTCCGCTGCTGGGAGCCCGCAACATCTCAAGGCG 6239
QY 6252 AGATTCCGGCGGGTACACACTGACCTTCGGGTGCCCCCGCAAGGTCCTCAGCGGCA 6311
DB 6240 AGATTCCGGCGGGTACACACTGACCTTCGGGTGCCCCCGCAAGGTCCTCAGCGGCA 6299
QY 6312 CGCGCTTCGTGGCGCGGAGTTCCTCTGGGTTCGGAGCTGGCGGAGCACATGGAGGCGCG 6371
DB 6300 CGCGCTTCGTGGCGCGGAGTTCCTCTGGGTTCGGAGCTGGCGGAGCACATGGAGGCGCG 6359
QY 6372 CTGGCTTCAGCTGCGCGCGGAGGCGCTGGCGCTGGCGCGCGCTTCTTGGAGAGCTG 6431
DB 6360 CTGGCTTCAGCTGCGCGCGGAGGCGCTGGCGCTGGCGCGCGCTTCTTGGAGAGCTG 6419
QY 6432 CGCGTGCACGCGCGAGAGCACGGCGTGGAGGACTTTTCCGTGAGCCAGCATGCTGGAG 6491

DB 6420 GCGGTGCACGCGCGACAGCAACGCGGTGGAGACTTTTCCGTGAGCCAGCATGCTGGAG 6479
QY 6492 GAGGTATTTCTGTACTTCTCCAAAGCACAGGGGAGAGGACAGAGACACCGAAGAGCAAG 6551
DB 6480 GAGGTATTTCTGTACTTCTCCAAAGCACAGGGGAGAGGACAGAGACACCGAAGAGCAAG 6539
QY 6552 GAGGCAAGGAGTGGAGTGGACCCCGCGCCAGGCGCTCAGCACCCCAAAACGCGTCAGCCAG 6611
DB 6540 GAGGCAAGGAGTGGAGTGGACCCCGCGCCAGGCGCTCAGCACCCCAAAACGCGTCAGCCAG 6599
QY 6612 TTCTCTCGATGACCTTACGACTGCGGAGACTGTCTCTGAGCCCTTCTCCCTCGCGGCGCC 6671
DB 6600 TTCTCTCGATGACCTTACGACTGCGGAGACTGTCTCTGAGCCCTTCTCCCTCGCGGCGCC 6659
QY 6672 GCGGGAGGCGCTTGGGAATGGCAAGGGCAAGGTAGAGTGTCTAGGAGCCCTTGACTCAGG 6731
DB 6660 GCGGGAGGCGCTTGGGAATGGCAAGGGCAAGGTAGAGTGTCTAGGAGCCCTTGACTCAGG 6719
QY 6732 CTGGCAGAGGCGCTGCTGCGCTGGAGAAATAAAGAGAGGCTGGAGAGAACCGCTGCTG 6791
DB 6720 CTGGCAGAGGCGCTGCTGCGCTGGAGAAATAAAGAGAGGCTGGAGAGAACCGCTGCTG 6779
QY 6792 GTGAAAAA 6803
DB 6780 GTGAAAAA 6791

RESULT 3
AX320362
LOCUS AX320362 6768 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent WO0187978.
ACCESSION AX320362
VERSION AX320362.1 GI:17901742
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Glucksmann, M. and Curtis, R. A.
TITLE 38594, a novel human transporter and uses thereof
JOURNAL Patent: WO 0187978-A 1 22-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..6768
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
165..6599
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD19391.1"
/db_xref="GI:17901743"
translation="MAFWTQLMLLWKNEWRRROPPLVLELLWPLFFLTVAVRHSH
PLEHHECHFPNKLPLSAGTVPNLOGLICNVANTCPOLTPGEEPRGLSNFNDSLVSR
LIADARTVLGSAHRTLAGLKLIALRAASTAQOPTKQSPFLPMLDVAELTSL
LURTESLALQAOGLPLHLEAGDLAEALLRSLRLEALLQRPSTGSPLELL
SEALCSVRGSPSTGPNLWYFASDLMEVLGPEPSALPDSSLPACSELIGLDSHP
LSRLWRLRLPLILGLKLPADPTFPRKLMAQVNRFTFELTLRLDRVEMWELGPRIF
TFMDSNNVAMLQRLQMDGRRORPRGCRDHMEALRFLDPGSGYSWQDAHADVG
HVLGTLGRVTEGLSLDKLEAAPSEALVSRALQLLAHRFAGWVFLGSDSDPTEH
PTDGLGHRVIRKIMDIDVTRTKIRDFRDPGPAADPLTDLRYVNGGFVYLQDLV
EAAVRLSGSANGRAGLIYQMPYPCYVDDVFLVLSRSLPLFLIAWISYVLTIVA
VYREKTRLDTRMAGLSRAVLWGLFSLCLGFLLSAALLVLKLDGIDILPYSPG
VVFPLFAFAVATVTFQSLISAFPSRANLAAACGLAFSLVPLVYLCVAMRDLRPA
GRVAASTLSPVAFGFCESLALLEEGEAGAHNVGTRTADVFSLAQVSGLLDLAA
LVGLATWLEAVCRGOYGIPEPNPFRESYVWCPPEPKSPAPCTPLPKVLEBAP
PLGSPGVSRLEKRFPGFPQALRGLSLDFYQGHITAFIGHSGAKTTLTSLGFL
PPSGSAFLIGHDVRSSMAAIRPHLGSCFQPNVFLDMLTVDHWFYGRUKGSAVW
GFQDRLQDLGVLSQSVQTRHLGSGMQRLSVAFVFGGQVVLDEPTAGDPAS
RGIWELLKRYREKTLILSTHLEDEALLGDRVAVAVAGRLCCCGSPFLRRHLGSG
YVLTLVKARLPLTTNEKADTDMEGSVDTROEKNGSQSGSRVGTGQLLALVQHWVPGAR

[illegible]

Db	2875	GGCGGCTGAAGGGTCTGAGTGC CGCTGTGTGTGGGCCCCGGAGCAGGACCGGCTCTACTGCAGG	2934
Qy	2986	ATGTGGGGCTGGTCTCTCCAAGCAGAGTGTGCAGACTCGCCACACTCTCTGTGGGATGCAAC	3045
Db	2935	ATGTGGGGCTGGTCTCCNAGCAGAGTGTGCAGACTCGCCACACTCTCTGTGTGGATGCNAC	2994
Qy	3046	GGAAAGCTGTCCGTGGGCAATTGCTTTGTGGCGGGCTCCCAAGTTGTATTCTCTGGACGAGC	3105
Db	2995	GGAAAGCTGTCCGTGGGCAATTGCTTTGTGGCGGGCTCCCAAGTTGTATTCTCTGGACGAGC	3054
Qy	3106	CTACGGCTGGCGTGGATCTGCTTCCGCGCGCGGTATTTGGGAGCTGCTCTCAATATACC	3165
Db	3055	CTACGGCTGGCGTGGATCTGCTTCCGCGCGCGGTATTTGGGAGCTGCTCTCAAAATACC	3114
Qy	3166	GAGAAAGTTCGACGCTGATCCTCTCCACCCACCACTGGATGAGGACGAGCTGCTGGGAG	3225
Db	3115	GAGAAAGTTCGACGCTGATCCTCTCCACCCACCACTGGATGAGGACGAGCTGCTGGGAG	3174
Qy	3226	ACCGTGTGGCTGTGTGGCAGGTGGCGCTTGCTGTCTGTGTGCTCCCCACTCTTCTCTGC	3285
Db	3175	ACCGTGTGGCGTGTGGCAGGTGGCGCTTGCTGTCTGTGTGCTCCCCACTCTTCTCTGC	3234
Qy	3286	GCCTGTACCTTGGGGTCCGGCTACTACCTGTACGCTGTGTGAAGGCGCGGCTGCCCTGACCA	3345
Db	3235	GCCTGTACCTTGGGGTCCGGCTACTACCTGTACGCTGTGTGAAGGCGCGGCTGCCCTGACCA	3294
Qy	3346	CCAATGAGAAGGCTGACACTGACATGAGAGGGCAGTGTGGACACACAGGACGAGAAAGAAGA	3405
Db	3295	CCAATGAGAAGGCTGACACTGACATGAGAGGGCAGTGTGGACACACAGGACGAGAAAGAAGA	3354
Qy	3406	ATGSCACCGAGGGCAGCAGAGTCCGGCACTCCTCAGCTGTGTGGCCCTTGSTAACAGCACTGGG	3465
Db	3355	ATGSCACCGAGGGCAGCAGAGTCCGGCACTCCTCAGCTGTGTGGCCCTTGSTAACAGCACTGGG	3414
Qy	3466	TGCCCCGGGCA CGGCTGTGTGGAGAGGTGGCAACAGAGCTGTGTGTGTGTGTGCCCTACA	3525
Db	3415	TGCCCGGGGCA CGGCTGTGTGGAGAGGTGGCAACAGAGCTGTGTGTGTGTGTGCCCTACA	3474
Qy	3526	CGGGTGCCCATGACGGCAGCTTCGGCACACTCTTCGAGAGCTAGACACGCGCTGGCGG	3585
Db	3475	CGGGTGCCCATGACGGCAGCTTCGGCACACTCTTCGAGAGCTAGACACGCGCTGGCGG	3534
Qy	3586	AGCTGAGGCTCACTTGGCTACGGGATCTCCGACACCAAGCCTCGAGGAGATCTTCTCTGAAGG	3645
Db	3535	AGCTGAGGCTCACTTGGCTACGGGATCTCCGACACCAAGCCTCGAGGAGATCTTCTCTGAAGG	3594
Qy	3646	TGTFGGAGGAGTGTGTGTGGCAACAGATATGAGAGATGTCAGCTGTGGGCGCAGCACTAT	3705
Db	3595	TGTFGGAGGAGTGTGTGTGGCAACAGATATGAGAGATGTCAGCTGTGGGCGCAGCACTAT	3654
Qy	3706	GCA CAGGCATTTGCTGGCCTAGAGCTAACCTGTGGGCTCAAGATGCCGCCACAGGAGACAG	3765
Db	3655	GCA CAGGCATTTGCTGGCCTAGAGCTAACCTGTGGGCTCAAGATGCCGCCACAGGAGACAG	3714
Qy	3766	CGCTGGAGAAACGGGAAACAGCTGGGTTCAGCCGCCAGAGACTGACCAGGGCTCTTGGGCCAG	3825
Db	3715	CGCTGGAGAAACGGGAAACAGCTGGGTTCAGCCGCCAGAGACTGACCAGGGCTCTTGGGCCAG	3774
Qy	3826	ACGCCGTGGCCCGGGTACAGGGCTGGGCACGTGACCCGCCACGAGCTCCAGGCCCTGTCTTC	3885
Db	3775	ACGCCGTGGCCCGGGTACAGGGCTGGGCACGTGACCCGCCACGAGCTCCAGGCCCTGTCTTC	3834
Qy	3886	TCAAGCGCTTTCTGCTTGC CGGCGGACGCGCGCGGCGCTGTTCGCCACAGATCGTCTGTC	3945
Db	3835	TCAAGCGCTTTCTGCTTGC CGGCGGACGCGCGCGGCGCTGTTCGCCACAGATCGTCTGTC	3894
Qy	3946	CTGGCCTCTTTGTGGGCTGGCCCTCGTGTTCAGGCTCACTGCTGCTCTTTTCGGGCACCT	4005
Db	3895	CTGGCCTCTTTGTGGGCTGGCCCTCGTGTTCAGGCTCACTGCTGCTCTTTTCGGGCACCT	3954
Qy	4006	ACCCGGCTCTCGGCTCAGTCCCAACATGTACGGTGTCTCAGGTGTCTTCTTTCAGTGAGG	4065
Db	3955	ACCCGGCTCTCGGCTCAGTCCCAACATGTACGGTGTCTCAGGTGTCTTCTTTCAGTGAGG	4014

QY 4066 ACGCCCCAGGGGACCTTGAGACGTGCCCGGCTGCTCGAGCGCTGCTGAGGAGGCGAGAC 4125
DB |||||
QY 4015 ACGCCCCAGGGGACCTTGAGACGTGCCCGGCTGCTCGAGCGCTGCTGAGGAGGCGAGAC 4074
DB |||||
QY 4126 TGGAGGAGCCCCAGTGGAGGAGTCCACAGAGTCTCGGACACAGAGTCTCTGCTG 4185
DB |||||
QY 4075 TGGAGGAGCCCCAGTGGAGGAGTCCACAGAGTCTCGGACACAGAGTCTCTGCTG 4134
DB |||||
QY 4186 AAGTGGCCCAAGGTCTTTGGCCAGTGGCAACTGAGACCCAGAGTCTCCATCCCAGGCTGCC 4245
DB |||||
QY 4135 AAGTGGCCCAAGGTCTTTGGCCAGTGGCAACTGAGACCCAGAGTCTCCATCCCAGGCTGCC 4194
DB |||||
QY 4246 AGTGTAGCCAGCCCCGGTGGCCCGCTGCTGCCCCGAGTGGCCCGCTGAGTGGTGC 4305
DB |||||
QY 4195 AGTGTAGCCCGCCCCGGTGGCCCGCTGCTGCCCCGAGTGGCCCGCTGAGTGGTGC 4254
DB |||||
QY 4306 CCCCTCCGCCCCAGGAGTGGCCCGCTGAGAGCTTACCCGCGCTGCTGGGAGTGGTTCAGAACCTGACAGGCCGGA 4365
DB |||||
QY 4255 CCCCTCCGCCCCAGGAGTGGCCCGCTGAGAGCTTACCCGCGCTGCTGGGAGTGGTTCAGAACCTGACAGGCCGGA 4314
DB |||||
QY 4366 ACCTGTCTGACTTCTCTGTGCTCAAGACCTTACCCGCGCTGCTGGGAGTGGTTCAGAACCTGACAGGCCGGA 4425
DB |||||
QY 4315 ACCTGTCTGACTTCTCTGTGCTCAAGACCTTACCCGCGCTGCTGGGAGTGGTTCAGAACCTGACAGGCCGGA 4374
DB |||||
QY 4426 AGAAGTGGGTGAATGAGGTCAAGGTACGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCC 4485
DB |||||
QY 4375 AGAAGTGGGTGAATGAGGTCAAGGTACGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCC 4434
DB |||||
QY 4486 TGCCCTCGGGCCAGAGTGGGGCCGCTCAGTGGAGGAGTGGTGGCGCTGCTGAGTCCCC 4545
DB |||||
QY 4435 TGCCCTCGGGCCAGAGTGGGGCCGCTCAGTGGAGGAGTGGTGGCGCTGCTGAGTCCCC 4494
DB |||||
QY 4546 TGCCCTCGGGGGCCCTCGACCGTGTCTGAAAGACCTCACAGCCTGGGCTCACAGCCTGG 4605
DB |||||
QY 4495 TGCCCTGGGGGGCCCTCGACCGTGTCTGAAAGACCTCACAGCCTGGGCTCACAGCCTGG 4554
DB |||||
QY 4606 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAGAGCTGGGACCTCCATGGTGGCCT 4665
DB |||||
QY 4555 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAGAGCTGGGACCTCCATGGTGGCCT 4614
DB |||||
QY 4666 TTGTCACCGGACGAGCAAGCAATCTCGTGTCTCAGCTGCGCCCGAGGCCGCGGCC 4725
DB |||||
QY 4615 TTGTCACCGGACGAGCAAGCAATCTCGTGTCTCAGCTGCGCCCGAGGCCGCGGCC 4674
DB |||||
QY 4726 ACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACAAAGCTGGGACCTCCATGGTGGCCTG 4785
DB |||||
QY 4675 ACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACAAAGCTGGGACCTCCATGGTGGCCTG 4734
DB |||||
QY 4786 AGGTGCACTGATGGCCCTCCTCGGTGGACGCTCCTGCTCCATCTGTGGTCTTTGGCA 4845
DB |||||
QY 4735 AGGTGCACTGATGGCCCTCCTCGGTGGACGCTCCTGCTCCATCTGTGGTCTTTGGCA 4794
DB |||||
QY 4846 TGTCTTTGTCGGGCGAGCTTCACTTTGTCCTCATTTGAGGAGGAGTCAACCGAGCCA 4905
DB |||||
QY 4795 TGTCTTTGTCGGGCGAGCTTCACTTTGTCCTCATTTGAGGAGGAGTCAACCGAGCCA 4854
DB |||||
QY 4906 AGCACCTGACGTCATGGGGGGCTGTCCCGCCACCTCTACTGAGTGGCAACTTCTCT 4965
DB |||||
QY 4855 AGCACCTGACGTCATGGGGGGCTGTCCCGCCACCTCTACTGAGTGGCAACTTCTCT 4914
DB |||||
QY 4966 GGGACATGTGTAACTACTTTGTTGGCCAGCATGCTGGTGTCTCATCTTTCTGGCCTTCC 5025
DB |||||
QY 4915 GGGACATGTGTAACTACTTTGTTGGCCAGCATGCTGGTGTCTCATCTTTCTGGCCTTCC 4974
DB |||||
QY 5026 AGCAGAGGGCATATGTGGCCCTCGCAACCTGCTGCTCTCTGCTGTGTGCTACTGT 5085
DB |||||
QY 4975 AGCAGAGGGCATATGTGGCCCTCGCAACCTGCTGCTCTCTGCTGTGTGCTACTGT 5034
DB |||||
QY 5086 ATGCTGTGTCATCACACCGCTCATGTATACCGAGCTCTCTCTCTGCTGCGGCCAGCA 5145
DB |||||
QY 5035 ATGCTGTGTCATCACACCGCTCATGTATACCGAGCTCTCTCTCTCTGCTGCGGCCAGCA 5094
DB |||||

QY 5146 CAGCCTATGTGGTGTCTACCTGCTCAATAAACCTCTTTATTGGCATCAATGGAAGCATGGCCA 5205
DB |||||
QY 5095 CAGCCTATGTGGTGTCTACCTGCTCAATAAACCTCTTTATTGGCATCAATGGAAGCATGGCCA 5154
DB |||||
QY 5206 CTTTGTGCTTTGAGCTCTTCTCTGATCAGAACTGCTCAGAGGTGAGCCGGATCTTGAAC 5265
DB |||||
QY 5155 CTTTGTGCTTTGAGCTCTTCTCTGATCAGAACTGCTCAGAGGTGAGCCGGATCTTGAAC 5214
DB |||||
QY 5266 AGGTCTTCTTATCTTTCCCACTTCTGCTTTGGGCCGGGGCTCATTTGACATGGTGGGA 5325
DB |||||
QY 5215 AGGTCTTCTTATCTTTCCCACTTCTGCTTTGGGCCGGGGCTCATTTGACATGGTGGGA 5274
DB |||||
QY 5326 ACCAGGCCCATGGCTGATGCTTTGAGCGCTTTGGGAGACAGGAGTTCAGTCAACCCCTG 5385
DB |||||
QY 5275 ACCAGGCCCATGGCTGATGCTTTGAGCGCTTTGGGAGACAGGAGTTCAGTCAACCCCTG 5334
DB |||||
QY 5386 GCTGGGAGGTGGTGGGAGAGAACCTTCTTGGCCATGGTGTATACAGGGGGCCCTCTTCTTCT 5445
DB |||||
QY 5335 GCTGGGAGGTGGTGGGAGAGAACCTTCTTGGCCATGGTGTATACAGGGGGCCCTCTTCTTCT 5394
DB |||||
QY 5446 TCTTTCACACTACTGCTGAGCACCGAAGCCAACTCTCTCCACAGCCAGGCGTGGTCTC 5505
DB |||||
QY 5395 TCTTTCACACTACTGCTGAGCACCGAAGCCAACTCTCTCCACAGCCAGGCGTGGTCTC 5454
DB |||||
QY 5506 TGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGGTCCAAG 5565
DB |||||
QY 5455 TGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGGTCCAAG 5514
DB |||||
QY 5566 GAGCCACCCAGGGGGATGTGTGTCTGAGAGAACTTGAACCAAGTATACCGTGGGGCAGA 5625
DB |||||
QY 5515 GAGCCACCCAGGGGGATGTGTGTCTGAGAGAACTTGAACCAAGTATACCGTGGGGCAGA 5574
DB |||||
QY 5626 GGATGCCAGCTGTGTGACCGCTGTGCTGGGGATTTCCCGCTGGTGGTGTGTTGGGCTG 5685
DB |||||
QY 5575 GGATGCCAGCTGTGTGACCGCTGTGCTGGGGATTTCCCGCTGGTGGTGTGTTGGGCTG 5634
DB |||||
QY 5686 TGGGTGTGAATGGAGCAGGGAGAACGCTCCACGTTTGCATGTGTGAGCGGGGACACATGG 5745
DB |||||
QY 5635 TGGGTGTGAATGGAGCAGGGAGAACGCTCCACGTTTGCATGTGTGAGCGGGGACACATGG 5694
DB |||||
QY 5746 CCAGCAGGGGGAGGCTGTGCTGAGGACCAAGGCTGGCCCGGGAACCCAGTGTGCTG 5805
DB |||||
QY 5695 CCAGCAGGGGGAGGCTGTGCTGAGGACCAAGGCTGGCCCGGGAACCCAGTGTGCTG 5754
DB |||||
QY 5806 ACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGAGCGGCCG 5865
DB |||||
QY 5755 ACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGAGCGGCCG 5814
DB |||||
QY 5866 AGCACCTGGAGCTGCTTTCGCGCTGCGCGGTGTCCCGAGGGCCAGGTTGCCAGACCG 5925
DB |||||
QY 5815 AGCACCTGGAGCTGCTTTCGCGCTGCGCGGTGTCCCGAGGGCCAGGTTGCCAGACCG 5874
DB |||||
QY 5926 CTGCTCAGGCTGCGCGCTGCGGACTCTCATGGTACGAGACGCGGCTGCGAGCACCT 5985
DB |||||
QY 5875 CTGCTCAGGCTGCGCGCTGCGGACTCTCATGGTACGAGACGCGGCTGCGAGCACCT 5934
DB |||||
QY 5986 ACAGCGAGGAGAACAAACGCAAGCTGCGAGCGCCCTGGCGCTGTGGGGACCCAGCG 6045
DB |||||
QY 5935 ACAGCGAGGAGAACAAACGCAAGCTGCGAGCGCCCTGGCGCTGTGGGGACCCAGCG 5994
DB |||||
QY 6046 TGGTGTCTTCTGGAGCAGCCGACCAAGGAGTGGACCCAGCGCGCGGCTTCTTTGGA 6105
DB |||||
QY 5995 TGGTGTCTTCTGGAGCAGCCGACCAAGGAGTGGACCCAGCGCGCGGCTTCTTTGGA 6054
DB |||||
QY 6106 ACAGCCTTTTGGCGCTGCGGAGGCGGCTTCAAGTGTGATGCTCACCTCCCATAGCATGG 6165
DB |||||
QY 6055 ACAGCCTTTTGGCGCTGCGGAGGCGGCTTCAAGTGTGATGCTCACCTCCCATAGCATGG 6114
DB |||||
QY 6166 AGGAGTGTGAAGCGCTCTGCTGCGCGCTAGCCATCATGTGTGAATGGGCGGTTCCGCTGCC 6225
DB |||||
QY 6115 AGGAGTGTGAAGCGCTCTGCTGCGCGCTAGCCATCATGTGTGAATGGGCGGTTCCGCTGCC 6174
DB |||||
QY 6226 TGGGAGCCCGGCAACATCTCAAGGGGAGATTTCCGGGGGGTCCACACACTGACCTGCGGG 6285
DB |||||

Db 6175 TGGGACCCCGCAACATCTCAAGGCGAGATTGGCGGGGTTCACACACTGACCCCTGGCGG 6234
Qy 6286 TGCCCGCCGCAAGTGTCCAGCCGCGAGCGGCTTTCGGGGCGCGAGATTCCTTGGGTCGG 6345
Db 6235 TGCCCGCCGCAAGTGTCCAGCCGCGAGCGGCTTTCGGGGCGCGAGATTCCTTGGGTCGG 6294
Qy 6346 AGCTGCGCGAGGACATGGAGGCGCTGCGCTTCCAGCTGCCGCGGGAGGCGCTGCG 6405
Db 6295 AGCTGCGCGAGGACATGGAGGCTGCGCTGCGCTTCCAGCTGCCGCGGGAGGCGCTGCG 6354
Qy 6406 CCTGGCGCGGCTTTCGGAGAGCTGGCGGTGCAACGCGCGAGAGCACGGCGTGGAGACT 6465
Db 6355 CCTGGCGCGGCTTTCGGAGAGCTGGCGGTGCAACGCGCGAGAGCACGGCGTGGAGACT 6414
Qy 6466 TTTCCTGAGCCAGACGATGCTGGAGGAGTATTTCTTGTACTTCTTCAAGGACAGGGGA 6525
Db 6415 TTTCCTGAGCCAGACGATGCTGGAGGAGTATTTCTTGTACTTCTTCAAGGACAGGGGA 6474
Qy 6526 AGGACGAGGACACCGAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCGAGGCC 6585
Db 6475 AGGACGAGGACACCGAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCGAGGCC 6534
Qy 6586 TGACGACACCCAAACGCGTCAGCCAGTTCCTCGATGACCCCTAGCACTGCCGAGACTGTGC 6645
Db 6535 TGACGACACCCAAACGCGTCAGCCAGTTCCTCGATGACCCCTAGCACTGCCGAGACTGTGC 6594
Qy 6646 TCTGAGCCTCCCTCCCTGCGGGGCGCGGGAGGCGCTGGGAATGCAAGGGCAAGGTA 6705
Db 6595 TCTGAGCCTCCCTCCCTGCGGGGCGCGGGAGGCGCTGGGAATGCAAGGGCAAGGTA 6654
Qy 6706 GAGTGCCCTAGAGCCCTGAGCTCAGGCTGGCAGAGGCGCTGCTGGAGAAATAA 6765
Db 6655 GAGTGCCCTAGAGCCCTGAGCTCAGGCTGGCAGAGGCGCTGCTGGAGAAATAA 6714
Qy 6766 GAGAGGCTGGAGAGGCGCTGCTGCTGGAATAAATAA 6804
Db 6715 GAGAGGCTGGAGAGGCGCTGCTGCTGGAATAAATAA 6753

RESULT 4
AX429481
LOCUS AX429481 7795 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 8 from Patent WO0234903.
ACCESSION AX429481
VERSION AX429481.1 GI:21540776
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Arnould-Reguigne, I., Deneffe, P., Chimini, G., Duverger, N.,
Osorio, Y.F., Prades, C. and Rosier, M.F.
TITLE Nucleic acid generating the abc7 gene, molecules modulating its
activity and therapeutic applications
JOURNAL Patent: WO 0234903-A 8 02-MAY-2002;
AVENTIS PHARMA SA (FR); INST NAT SANTE RECH MED (FR)
FEATURES
source 1. .7795
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.6%; Score 6710; DB 6; Length 7795;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6722; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 56 TGGGAGCCCTCGAAGAGCTTCCAGGAACCTTCGCTGTGGGATAAAGGAATGAGGTTTC 115
Db 1054 TGCCCTGCTCTAATCAGAGCTTCCAGGAACCTTCGCTGTGGGATAAAGGAATGAGGTTTC 1113

Qy 116 AGAAGGGGCGAGGAGTGTCCCGCAGCGCACCGCAGCTTTCAGGCCCGACCGCTGTGCT 175
Db 1114 AGAAGGGGCGAGGAGTGTCCCGCAGCGCACCGCAGCTTTCAGGCCCGACCGCTGTGCT 1173
Qy 176 GACCTCTCTGTCCCGTCCCGCCAGTCTCACCATGGGCTTCTGACACAGCTGATGCT 235
Db 1174 GACCTCTCTGTCCCGTCCCGTCCCGCCAGTCTCACCATGGGCTTCTGACACAGCTGATGCT 1233
Qy 236 GCTGCTCTGGAGAAATTTTATGTATCGCGGAGACAGCGGTCCAGCTCCTGCTCGAATT 295
Db 1234 GCTGCTCTGGAGAAATTTTATGTATCGCGGAGACAGCGGTCCAGCTCCTGCTCGAATT 1293
Qy 296 GCTGTGGCCT 355
Db 1294 GCTGTGGCCT 1353
Qy 356 GGAGCACCATGATGTCACATTTCCAAACAAAGCCACTGCCATCGCGGGCACCGTGCCTG 415
Db 1354 GGAGCACCATGATGTCACATTTCCCAAAACAAAGCCACTGCCATCGCGGGCACCGTGCCTG 1413
Qy 416 GCTCCAGGGTCTCATCTGTAAATGTGAACAAACACTGCTTTCCGAGCTGACACCGGCGA 475
Db 1414 GCTCCAGGGTCTCATCTGTAAATGTGAACAAACACTGCTTTCCGAGCTGACACCGGCGA 1473
Qy 476 GGAGCCCGGGCGCTGAGCAACTTTCAACGACTCCCTGGTCTCCCGGCTGTAGCCGATGC 535
Db 1474 GGAGCCCGGGCGCTGAGCAACTTTCAACGACTCCCTGGTCTCCCGGCTGTAGCCGATGC 1533
Qy 536 CCGCACTGTCTGGGAGGGCCAGTCCCAAGGAACGCTGGCTGAGGAGCTGAT 595
Db 1534 CCGCACTGTCTGGGAGGGCCAGTCCCAAGGAACGCTGGCTGAGGAGCTGAT 1593
Qy 596 CGCCACGCTGAGGGCTGACAGGAGCAGCGCCAGCTCAACCAACCAAGCAGTCTCCACT 655
Db 1594 CGCCACGCTGAGGGCTGACAGGAGCAGCGCCAGCTCAACCAACCAAGCAGTCTCCACT 1653
Qy 656 GGAACCAACCATGCTGAGTGTGCGGAGCTGTGAGCTCACTGCTGCGACGGAATCCCT 715
Db 1654 GGAACCAACCATGCTGAGTGTGCGGAGCTGTGAGCTCACTGCTGCGACGGAATCCCT 1713
Qy 716 GGGGTTGGCACTGGGCGCAAGCCAGAGCCCTTTGCAACAGCTTTTGAGGCGCTGAGGA 775
Db 1714 GGGGTTGGCACTGGGCGCAAGCCAGAGCCCTTTGCAACAGCTTTTGAGGCGCTGAGGA 1773
Qy 776 CTTGGCCCGAGGAGCTCTGCGGCTGCGAGCTGTGGTGGAGCTTTCGGCAGCTGTCAGAG 835
Db 1774 CTTGGCCCGAGGAGCTCTGCGGCTGCGAGCTGTGGTGGAGCTTTCGGCAGCTGTCAGAG 1833
Qy 836 ACCCCGAGGAGCAGCGGGCCCCCTGGAGTGTGCTGTCAGAGGCCCTCTGCAAGTGTCAAGGG 895
Db 1834 ACCCCGAGGAGCAGCGGGCCCCCTGGAGTGTGCTGTCAGAGGCCCTCTGCAAGTGTCAAGGG 1893
Qy 896 ACCTAGCAGCAGTGGGCCCCCTCCCTCAACTGGTACGAGGCTAGTGAACCTGATGAGCT 955
Db 1894 ACCTAGCAGCAGTGGGCCCCCTCCCTCAACTGGTACGAGGCTAGTGAACCTGATGAGCT 1953
Qy 956 GGTGGGGCAGGAGCAGAAATCCGCTTCCAGACAGCAGCTGAGCCCCCGCTGCTCGGA 1015
Db 1954 GGTGGGGCAGGAGCAGAAATCCGCTTCCAGACAGCAGCTGAGCCCCCGCTGCTCGGA 2013
Qy 1016 GCTGATTGGAGCCCTGAGACAGCCAGCCGCTGCTGCCCTGCTCTGGAGAGCCCTGAAGCC 1075
Db 2014 GCTGATTGGAGCCCTGAGACAGCCAGCCGCTGCTGCCCTGCTCTGGAGAGCCCTGAAGCC 2073
Qy 1076 TCTGATCCTCGGAAGCTACTCTTTTGCACAGATACACCTTTTACCCGGAAGCTCATGGC 1135
Db 2074 TCTGATCCTCGGAAGCTACTCTTTTGCACAGATACACCTTTTACCCGGAAGCTCATGGC 2133
Qy 1136 CAGGTGTAACCGGACCTTCAGAGAGCTCACCTGCTGAGGAGTGTCCGGAGGCTGTGGGA 1195
Db 2134 CAGGTGTAACCGGACCTTCAGAGAGCTCACCTGCTGAGGAGTGTCCGGAGGCTGTGGGA 2193
Qy 1196 GATGCTGGGAGCCCGGATCTTCACTTTCATGAACGACAGTTTCCAATGTGGCCATGCTGCA 1255

Db 2194 GATGCTGGAGCCCCGGATCTTCACTTTCATGAACACAGATTCCAAATGTGGCCATGCTGCA 2253
Qy 1256 GCGGCTCTGCAAGATGCAAGGATGAAGGAAGAGGAGCCAGACCTGGAGGCGGGGACCA 1315
Db 2254 GCGGCTCTGCAAGATGCAAGGATGAAGGAAGAGGAGCCAGACCTGGAGGCGGGGACCA 2313
Qy 1316 CATGAGGCGCTGCGATCTTCTTGAGACCTGGAGAGCGGTGGCTACAGCTGGCAGGAGCG 1375
Db 2314 CATGAGGCGCTGCGATCTTCTTGAGACCTGGAGAGCGGTGGCTACAGCTGGCAGGAGCG 2373
Qy 1376 ACACGCTGATGTGGGCGCACTGTGTGGGACGCTGGGCGGAGGTGACGGAGTGCCTGCTCTT 1435
Db 2374 ACACGCTGATGTGGGCGCACTGTGTGGGACGCTGGGCGGAGGTGACGGAGTGCCTGCTCTT 2433
Qy 1436 GGACAAGCTGGAGGCGGCACTCTCAGAGGCGAGCCCTGGTGTGGGCGGCTGCAACTGCT 1495
Db 2434 GGACAAGCTGGAGGCGGCACTCTCAGAGGCGAGCCCTGGTGTGGGCGGCTGCAACTGCT 2493
Qy 1496 GCGGAAACATGATTTCTGGGCGGCGTCTGCTCTTTGGGACCTGAGGACTCTTACAGACC 1555
Db 2494 GCGGAAACATGATTTCTGGGCGGCGTCTGCTCTTTGGGACCTGAGGACTCTTACAGACC 2553
Qy 1556 CACAGACACCAACCCAGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1615
Db 2554 CACAGACACCAACCCAGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2613
Qy 1616 CATTGACGTGTGACGAGGACCAATAGATCAGGAGCAGGTTTGGGACCTGGGCGGCGGCGGCGG 1675
Db 2614 CATTGACGTGTGACGAGGACCAATAGATCAGGAGCAGGTTTGGGACCTGGGCGGCGGCGGCGG 2673
Qy 1676 GCGGAAACCCCTGACCGCTGCGCTACGTGTGGGCGGCTTCTGTATACCTGCAAGACCT 1735
Db 2674 GCGGAAACCCCTGACCGCTGCGCTACGTGTGGGCGGCTTCTGTATACCTGCAAGACCT 2733
Qy 1736 GGTGGAGCGTGACCGCTGCGGCTGCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1795
Db 2734 GGTGGAGCGTGACCGCTGCGGCTGCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2793
Qy 1796 GCAGCAGATGCCCTATCCGCTGCTATGTGGAACGAGTGTCTGCGTGTGCTGAGCGGCTC 1855
Db 2794 GCAGCAGATGCCCTATCCGCTGCTATGTGGAACGAGTGTCTGCGTGTGCTGAGCGGCTC 2853
Qy 1856 GCTGCGCTCTTCTGACGCTGGCTGGATCTACTCTCGTGAACCTGACAGTGAAGGCGGT 1915
Db 2854 GCTGCGCTCTTCTGACGCTGGCTGGATCTACTCTCGTGAACCTGACAGTGAAGGCGGT 2913
Qy 1916 GGTGGGAGAGGAGAGCGGCTGCGGGAACCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1975
Db 2914 GGTGGGAGAGGAGAGCGGCTGCGGGAACCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2973
Qy 1976 GGTGCTCTGGCTAGGCTGGTTCTCAGCTGCTCGGCGGCTTCTGCTCAGCGCGGCT 2035
Db 2974 GGTGCTCTGGCTAGGCTGGTTCTCAGCTGCTCGGCGGCTTCTGCTCAGCGCGGCT 3033
Qy 2036 GCTGCTCTGGCTCAGCTGGGGAACATCTCTCCCTACAGCAACCGGCGGCTGGTCTT 2095
Db 3034 GCTGCTCTGGCTCAGCTGGGGAACATCTCTCCCTACAGCAACCGGCGGCTGGTCTT 3093
Qy 2096 CTTGTTCTGGCAGCTTTCGGGTGGGCAACCGGTGACCCAGAGCTTCTGCTCAGCGGCTT 2155
Db 3094 CTTGTTCTGGCAGCTTTCGGGTGGGCAACCGGTGACCCAGAGCTTCTGCTCAGCGGCTT 3153
Qy 2156 CTTCTCCGCGCAACCTGGCTGGGCTGGGCGGCTGGGCGGCTTCTCTCCCTCTACCT 2215
Db 3154 CTTCTCCGCGCAACCTGGCTGGGCTGGGCGGCTGGGCGGCTTCTCTCCCTCTACCT 3213
Qy 2216 GCCCTACGTGTGTGTGGCTTGGCGGGAACCGGCTGCGCGGCTGGGCGGCTGGGCGGCT 2275
Db 3214 GCCCTACGTGTGTGTGGCTTGGCGGGAACCGGCTGCGCGGCTGGGCGGCTGGGCGGCT 3273
Qy 2276 GAGCTGTGTGCGCGGCTTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2335

Db 3274 GAGCTGTGTGCTGCGCGGCTTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 3333
Qy 2336 GCAGGCGAGGCGCGAGTGGGCACAACTGTGGGCAACCGGCGCTACGCGAGAGCTTCTCAG 2395
Db 3334 GCAGGCGAGGCGCGAGTGGGCACAACTGTGGGCAACCGGCGCTACGCGAGAGCTTCTCAG 3393
Qy 2396 CTTGCGCGAGTCTTCTGCGCTTCTGCTGTGGAGCGGCGGCTTACGGGCTTGGGCGGCT 2455
Db 3394 CTTGCGCGAGTCTTCTGCGCTTCTGCTGTGGAGCGGCGGCTTACGGGCTTGGGCGGCT 3453
Qy 2456 GTACTGTGAAGCTGTGTGGGCGGAGTGGGATCCCTGAACTGAAATTTTCTCTT 2515
Db 3454 GTACTGTGAAGCTGTGTGGGCGGAGTGGGATCCCTGAACTGAAATTTTCTCTT 3513
Qy 2516 TCGGAGGAGTACTGTGTGGGCGGAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2575
Db 3514 TCGGAGGAGTACTGTGTGGGCGGAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3573
Qy 2576 GCTGGACCCAAAGGTGTGTGAAGAGGCAACCGGCGGCGGCTGAGTCTGCGGTATCTCGT 2635
Db 3574 GCTGGACCCAAAGGTGTGTGAAGAGGCAACCGGCGGCGGCTGAGTCTGCGGTATCTCGT 3633
Qy 2636 TCGCAGCTGTGAGAGGCGCTTCTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2695
Db 3634 TCGCAGCTGTGAGAGGCGCTTCTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 3693
Qy 2696 GGACTTCTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2755
Db 3694 GGACTTCTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3753
Qy 2756 CACCTGTTCATTTGAGTGGGCTTCTCCACCCAGTGGTGGCTTCTGCTTCTCTGCGG 2815
Db 3754 CACCTGTTCATTTGAGTGGGCTTCTCCACCCAGTGGTGGCTTCTGCTTCTCTGCGG 3813
Qy 2816 CCAGAGCTCGGCTCAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2875
Db 3814 CCAGAGCTCGGCTCAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 3873
Qy 2876 CAACTGCTGTTTGACATGTGTGAGCGGAGCAGTGTGTTCTATGGGCGGCTCAA 2935
Db 3874 CAACTGCTGTTTGACATGTGTGAGCGGAGCAGTGTGTTCTATGGGCGGCTGAA 3933
Qy 2936 GGGTCTGAGTGGCGCTGTAGTGGGCGGCGGAGCAGTGTGCTGCAAGATGTGGGCT 2995
Db 3934 GGGTCTGAGTGGCGCTGTAGTGGGCGGCGGAGCAGTGTGCTGCAAGATGTGGGCT 3993
Qy 2996 GGTCTCAAAGCAGAGTGTGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 3055
Db 3994 GGTCTCAAAGCAGAGTGTGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 4053
Qy 3056 CGTGGCGGCTTGTGGGCGGCTCCCAAGTGTGTTATCTGAGCAGGCTACGGCTGG 3115
Db 4054 CGTGGCGGCTTGTGGGCGGCTCCCAAGTGTGTTATCTGAGCAGGCTACGGCTGG 4113
Qy 3116 CGTGGATCTGCTTCCGCGCGGCTATTTGGGAGCTGCTGCTCAAATACCGAGAGGCTCG 3175
Db 4114 CGTGGATCTGCTTCCGCGCGGCTATTTGGGAGCTGCTGCTCAAATACCGAGAGGCTCG 4173
Qy 3176 CAGCTGATCTCTCAGCCACCACTGGATGAGGAGAGTGTGCGGAGACCGGTGGG 3235
Db 4174 CAGCTGATCTCTCAGCCACCACTGGATGAGGAGAGTGTGCGGAGACCGGTGGG 4233
Qy 3236 TGTGGTGGCAGGTGGGCGGCTTGTGCTGTGGGCTCCCACTTCTGCGGCGGCTACCT 3295
Db 4234 TGTGGTGGCAGGTGGGCGGCTTGTGCTGTGGGCTCCCACTTCTGCGGCGGCTACCT 4293
Qy 3296 GGGTCTCGGCTACTACTGCTGAGCGTGGTGAAGGCGGCGGCTGCGGCTGAGGAA 3355
Db 4294 GGGTCTCGGCTACTACTGCTGAGCGTGGTGAAGGCGGCGGCTGCGGCTGAGGAA 4353
Qy 3356 GGTGACACTGACATGAGGCGGAGTGTGGACACAGGCGAGGAAAGAAAGAAATGCGAGCA 3415
Db 4354 GGTGACACTGACATGAGGCGGAGTGTGGACACAGGCGAGGAAAGAAAGAAATGCGAGCA 4413

QY 3416 GGCAGCAGAGTGGGCACTCTCAGCTGCTGGCCCTGGTACAGCACTGGGTGCCCGGGC 3475
DB |||||
QY 4414 GGGCAGCAGAGTCGGCACTCTCAGCTGCTGGCCCTGGTACAGCACTGGGTGCCCGGGC 4473
DB |||||
QY 3476 ACGGCTGGTGGAGAGTGGCCACACAGAGCTGGTGGTCTGCTGCCCTACACGGGTGCCCA 3535
DB |||||
QY 4474 ACGGCTGGTGGAGAGTGGCCACACAGAGCTGGTGGTCTGCCCTACACGGGTGCCCA 4533
DB |||||
QY 3536 TGACGGAGCTTGGCCACACTCTTCGAGAGCTAGACACCGGGTGGCGAGCTGAGGCT 3595
DB |||||
QY 4534 TGACGGAGCTTGGCCACACTCTTCGAGAGCTAGACACCGGGTGGCGAGCTGAGGCT 4593
DB |||||
QY 3596 CACTGGCTACGGGATCTCCGACACAGCTTCGAGGAGATCTTCTGAAGTGGTGGAGGA 3655
DB |||||
QY 4594 CACTGGCTACGGGATCTCCGACACAGCTTCGAGGAGATCTTCTGAAGTGGTGGAGGA 4653
DB |||||
QY 3656 GTGTGCTGGGACACAGATATGGAGATGGCAGCTGGGGCAGCACCTATGCA CAGGCAT 3715
DB |||||
QY 4654 GTGTGCTGGGACACAGATATGGAGATGGCAGCTGGGGCAGCACCTATGCA CAGGCAT 4713
DB |||||
QY 3716 TGCTGGCTAGACGTAAACCTTGGGCTCAAGATGGCCCA CAGAGACAGCGCTGGAGAA 3775
DB |||||
QY 4714 TGCTGGCTAGACGTAAACCTTGGGCTCAAGATGGCCCA CAGAGACAGCGCTGGAGAA 4773
DB |||||
QY 3776 CGGGAACAGCTGGGTACGCCACAGAGATGACACAGGGCTCTGGGCCAGACGCCGTGGG 3835
DB |||||
QY 4774 CGGGAACAGCTGGGTACGCCACAGAGATGACACAGGGCTCTGGGGCAGACGCCGTGGG 4833
DB |||||
QY 3836 CGGGTACAGGGCTGGGCACTGACCCGCCAGAGCTCCAGGCCCTGTCTTCAAGGGCTT 3895
DB |||||
QY 4834 CGGGTACAGGGCTGGGCACTGACCCGCCAGAGCTCCAGGCCCTGCTTCTCAAGGGCTT 4893
DB |||||
QY 3896 TCTGCTTGGCCGCGCAGCGCCGCGCTGTGGCCAGATCGTGTGCTGCTGCCCTTT 3955
DB |||||
QY 4894 TCTGCTTGGCCGCGCAGCGCCGCGCTGTGGCCAGATCGTGTGCTGCTGCCCTTT 4953
DB |||||
QY 3956 TGTGGGCTGGCCCTGTGTTCAGCTCATCGTGCCTCTTTCGGGCACTACCGGCTCT 4015
DB |||||
QY 4954 TGTGGGCTGGCCCTGTGTTCAGCTCATCGTGCCTCTTTCGGGCACTACCGGCTCT 5013
DB |||||
QY 4016 GCGGCTCAGTCCCAACCATGTACGGTGTCTCAGGTGTCTTCTTTCAGTGAGGACGCCCCAGG 4075
DB |||||
QY 5014 GCGGCTCAGTCCCAACCATGTACGGTGTCTCAGGTGTCTTCTTTCAGTGAGGACGCCCCAGG 5073
DB |||||
QY 4076 GGAACCTGGACGTGCCCGCTGTCTGAGGCGCTGTGACGAGGCGCAGGAATGGAGAGCC 4135
DB |||||
QY 5074 GGAACCTGGACGTGCCCGCTGTCTGAGGCGCTGTGACGAGGCGCAGGAATGGAGGAGCC 5133
DB |||||
QY 4136 CCAGTGACATAGCTCCCA CAGGTTCTCGGCACCAAGAGTTCTCTGCTGAAGTGGCCAA 4195
DB |||||
QY 5134 CCAGTGACATAGCTCCCA CAGGTTCTCGGCACCAAGAGTTCTCTGCTGAAGTGGCCAA 5193
DB |||||
QY 4196 GGTCTTGGCCAGTGGCAACTGGACCC CAGAGTCTCCATCCCA CAGCTGCCAGTGTAGCCA 4255
DB |||||
QY 5194 GGTCTTGGCCAGTGGCAACTGGACCC CAGAGTCTCCATCCCA CAGCTGCCAGTGTAGCCG 5253
DB |||||
QY 4256 GCCCGGTGCCCGGCGCTGTGCCGACTCCCGGCTGACGTGTGGTCCCCCTCGGC 4315
DB |||||
QY 5254 GCCCGGTGCCCGGCGCTGTGCCGACTCCCGGCTGACGTGTGGTCCCCCTCGGC 5313
DB |||||
QY 4316 CCAGGAGTACCGGCTCTGGGAGTGGTTCAAGACCTTACAGGCCCGGAACCTGTCTGA 4375
DB |||||
QY 5314 CCAGGAGTACCGGCTCTGGGAGTGGTTCAAGACCTTACAGGCCCGGAACCTGTCTGA 5373
DB |||||
QY 4376 CTTCTGGTCAAGACCTTACCGGCTGTGGCCAGGGCTTGAAGACTAAGAGTGGGT 4435
DB |||||
QY 5374 CTTCTGGTCAAGACCTTACCGGCTGTGGCCAGGGCTTGAAGACTAAGAGTGGGT 5433
DB |||||
QY 4436 GAATGAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCTGCTCGCTGGG 4495
DB |||||
QY 5434 GAATGAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCTGCTCGCTGGG 5493

QY 4496 CCAAGAGTTGGGCGGCTCAGTGGAGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCGCGG 4555
DB |||||
QY 5494 CCAAGAGTTGGGCGGCTCAGTGGAGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCGCGG 5553
DB |||||
QY 4556 GGCCCTCGACCGGTGCTCTGAAAACTCACAGCCTGGGTTCACAGCCTGGATGCTCAGGA 4615
DB |||||
QY 5554 GGCCCTCGACCGGTGCTCTGAAAACTCACAGCCTGGGTTCACAGCCTGGATGCTCAGGA 5613
DB |||||
QY 4616 CAGTCTCAAGATCTGGTTCAAACAAAGCTGGCACTCCATGGTGGCCTTGTTCACCG 4675
DB |||||
QY 5614 CAGTCTCAAGATCTGGTTCAAACAAAGCTGGCACTCCATGGTGGCCTTGTTCACCG 5673
DB |||||
QY 4676 AGCCAGCAACGCAATCTCCGTCTCACCTGCCCGCCAGGCGCCGCGCCACACAG 4735
DB |||||
QY 5674 AGCCAGCAACGCAATCTCCGTCTCACCTGCCCGCCAGGCGCCGCGCCACACAG 5733
DB |||||
QY 4736 CATCACCACTCAACCAACCCCTTGAACTCAACAAAGGAGCAGCTGTCTGAGGCTGCACT 4795
DB |||||
QY 5734 CATCACCACTCAACCAACCCCTTGAACTCAACAAAGGAGCAGCTGTCTGAGGCTGCACT 5793
DB |||||
QY 4796 GATGGCTCTCTCGGTGGAGCTCTCTCTCTCCATCTGTGTGGTCTTGTGCCATGCTTGT 4855
DB |||||
QY 5794 GATGGCTCTCTCGGTGGAGCTCTCTCTCTCCATCTGTGTGGTCTTGTGCCATGCTTGT 5853
DB |||||
QY 4856 CCGGCGCAGCTTCACTCTTGTCTCATTTGAGAGCGAGTCAACCCGAGCCAAAGCACCTGCA 4915
DB |||||
QY 5854 CCGGCGCAGCTTCACTCTTGTCTCATTTGAGAGCGAGTCAACCCGAGCCAAAGCACCTGCA 5913
DB |||||
QY 4916 GCTCATGGGGGGCTGTGCCCAACCTCTACTGGCTTGGCAACTTCTCTGGGACATGTG 4975
DB |||||
QY 5914 GCTCATGGGGGGCTGTGCCCAACCTCTACTGGCTTGGCAACTTCTCTGGGACATGTG 5973
DB |||||
QY 4976 TAACTACTTGGTGGCAGGATGCACTGTGTGCTCATCTTCTGGCCTTCCAGCAGAGGGC 5035
DB |||||
QY 5974 TAACTACTTGGTGGCAGGATGCACTGTGTGCTCATCTTCTGGCCTTCCAGCAGAGGGC 6033
DB |||||
QY 5036 ATATGTGGCCCTGCCAACTGCTGCTCTCTGCTGTGTGCTACTACTACTGTATGGCTGGTC 5095
DB |||||
QY 6034 ATATGTGGCCCTGCCAACTGCTGCTCTCTGCTGTGTGCTACTACTACTGTATGGCTGGTC 6093
DB |||||
QY 5096 GATCACACGCTCATGTATACCCAGCTCTCTTCTTCTCCGTGCCCAGACAGCCTATGT 5155
DB |||||
QY 6094 GATCACACGCTCATGTATACCCAGCTCTCTTCTTCTCCGTGCCCAGACAGCCTATGT 6153
DB |||||
QY 5156 GGTGCTCACTGCATATAA CTTCTTTATGGCATCAATGGAGCATGGCCACCTTGTGCT 5215
DB |||||
QY 6154 GGTGCTCACTGCATATAA CTTCTTTATGGCATCAATGGAGCATGGCCACCTTGTGCT 6213
DB |||||
QY 5216 TGAGCTCTTCTCTGATCAGAGCTGCAGAGGTGAGCCGATCTTGAACAGAGTCTTCTCT 5275
DB |||||
QY 6214 TGAGCTCTTCTCTGATCAGAGCTGCAGAGGTGAGCCGATCTTGAACAGAGTCTTCTCT 6273
DB |||||
QY 5276 TATCTTCCCCCACTTCTGCTTGGCCCGGGGCTCATTTGACATGGTTCGGAAACAGGCCAT 5335
DB |||||
QY 6274 TATCTTCCCCCACTTCTGCTTGGCCCGGGGCTCATTTGACATGGTTCGGAAACAGGCCAT 6333
DB |||||
QY 5336 GGTGATGCTTGTAGCGCTTGGGACACAGGCAAGTTCAGTCAACCCCTGCGCTGGGAGGT 5395
DB |||||
QY 6334 GGTGATGCTTGTAGCGCTTGGGACACAGGCAAGTTCAGTCAACCCCTGCGCTGGGAGGT 6393
DB |||||
QY 5396 GGTGCGCAAGAACCTCTTGGCCATGTGTATACAGGGGGCCCTCTTCTCTCTTCTCACT 5455
DB |||||
QY 6394 GGTGCGCAAGAACCTCTTGGCCATGTGTATACAGGGGGCCCTCTTCTCTCTTCTCACT 6453
DB |||||
QY 5456 ACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAGCCAGGGTGAAGTCTCTGCACTCTCT 5515
DB |||||
QY 6454 ACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAGCCAGGGTGAAGTCTCTGCACTCTCT 6513
DB |||||
QY 5516 GGGAGAGGAGCAGAGATGTAGCCCTGAA CCGGAGCGGGTGGTCCAGGAGCCACCCA 5575
DB |||||
QY 6514 GGGAGAGGAGCAGAGATGTAGCCCTGAA CCGGAGCGGGTGGTCCAGGAGCCACCCA 6573
DB |||||
QY 5576 GGGGATGTGTGGTCTCTGAGGAACCTTGACAGGTATACCGTGGGCGAGAGGATGCCAGC 5635

Db 6574 GGGGATGTTGGTGTGAGAACTTGACCAAGGTATACCTGGCGAGAGATCCACG 6633
Qy 5636 TGTGACCGCTTGTGCTGGGATTCCTCCCTGGTGAAGTGTGTTGGGCTGCTGGGTGCA 5695
Db 6634 TGTGACCGCTTGTGCTGGGATTCCTCCCTGGTGAAGTGTGTTGGGCTGCTGGGTGCA 6693
Qy 5696 TGGACAGGGAAGAGCTGCATCTTTGCGATGTGTGACGGGGACACATTTGGCCAGCAGGG 5755
Db 6694 TGGACAGGGAAGAGCTGCATCTTTGCGATGTGTGACGGGGACACATTTGGCCAGCAGGG 6753
Qy 5756 CGAGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAACCCAGTGTGCGCACCTCAGCAT 5815
Db 6754 CGAGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAACCCAGTGTGCGCACCTCAGCAT 6813
Qy 5816 GGGATACTGCTCAATCCGATGCCATCTTTGAGCTGTGTGACGGCCCGGAGCACCTGGA 5875
Db 6814 GGGATACTGCTCAATCCGATGCCATCTTTGAGCTGTGTGACGGCCCGGAGCACCTGGA 6873
Qy 5876 GCTGCTGGCGCTGCGCGGTGCTCCGAGGCCACAGTGTGCCAGACCGTGGCTCAGG 5935
Db 6874 GCTGCTGGCGCTGCGCGGTGCTCCGAGGCCACAGTGTGCCAGACCGTGGCTCAGG 6933
Qy 5936 CTTGCGCGCTGCGGACTCTCATGTATCGCAGACCGGCTGCGAGCACCTACAGCGAGG 5995
Db 6934 CTTGCGCGCTGCGGACTCTCATGTATCGCAGACCGGCTGCGAGCACCTACAGCGAGG 6993
Qy 5996 GAACAAACGCAAGCTGCGCGCTGCGCGTGTGGGACCGGCTGGTGTCT 6055
Db 6994 GAACAAACGCAAGCTGCGCGCTGCGCGTGTGGGACCGGCTGGTGTCT 7053
Qy 6056 GGAAGCGCGACACAGGATCGACCCAGCGCGCGCTTCTTTTGGAAACAGCCTTTT 6115
Db 7054 GGAAGCGCGACACAGGATCGACCCAGCGCGCGCTTCTTTTGGAAACAGCCTTTT 7113
Qy 6116 GCGCTGTGCGGAGGCGCTTCACTGATCTCATCTCCATGATGAGGAGTGTGA 6175
Db 7114 GCGCTGTGCGGAGGCGCTTCACTGATCTCATCTCCATGATGAGGAGTGTGA 7173
Qy 6176 AGCGCTGTCTGCGCCCTAGCATCATGTGTAATGGCGGTTCGCTGCTGCGGCGGCC 6235
Db 7174 AGCGCTGTCTGCGCCCTAGCATCATGTGTAATGGCGGTTCGCTGCTGCGGCGGCC 7233
Qy 6236 GCAACATCTCAAGGCGAGATTGCGGCGGGTCACTGACCTCTGCGGCTGCGGCGCGC 6295
Db 7234 GCAACATCTCAAGGCGAGATTGCGGCGGGTCACTGACCTCTGCGGCTGCGGCGCGC 7293
Qy 6296 AAGTCCCGAGCGGCGCGCTTCTGCGGCGCGAGTTCCTGCGTGGAGCTGCGCGA 6355
Db 7294 AAGTCCCGAGCGGCGCGCTTCTGCGGCGCGAGTTCCTGCGTGGAGCTGCGCGA 7353
Qy 6356 GGACATGGAGCGCGCTTCTGCGCTTCCAGCTGCGCGGCGGCGCTGCGCCCTGGCGCG 6415
Db 7354 GGACATGGAGCGCGCTTCTGCGCTTCCAGCTGCGCGGCGGCGCTGCGCCCTGGCGCG 7413
Qy 6416 CGTCTTTGGAGAGCTGCGGCTGCAAGCGCGAGAGCACGCGCTGGAGGACTTTTCGCTGAG 6475
Db 7414 CGTCTTTGGAGAGCTGCGGCTGCAAGCGCGAGAGCACGCGCTGGAGGACTTTTCGCTGAG 7473
Qy 6476 CCAGACGATGCTGAGGAGGATTTCTTGTATCTTCTCAAGGACCAAGGGAAGGACGAGGA 6535
Db 7474 CCAGACGATGCTGAGGAGGATTTCTTGTATCTTCTCAAGGACCAAGGGAAGGACGAGGA 7533
Qy 6536 CACCGAAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCCAGCGCTGCGACGCC 6595
Db 7534 CACCGAAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCCAGCGCTGCGACGCC 7593
Qy 6596 CAACCGCTCAGCCAGTTCCTCGATGACCTTAGACCTGCGAGACTGTGCTCTGAGCCTC 6655
Db 7594 CAACCGCTCAGCCAGTTCCTCGATGACCTTAGACCTGCGAGACTGTGCTCTGAGCCTC 7653
Qy 6656 CCTCCCTGCGGCGCGCGGAGGCCCTGGGAATGGCAAGGCAAGGTAGAGTGCCTAG 6715

Db 7654 CCTCCCTGCGGCGCGCGGAGCCCTGGGAATGGCAAGGCAAGGTAGAGTGCCTAG 7713
Qy 6716 GAGCCTCGACTCAGGCTGGCAGAGGGGCTGCTGCTGGCAAAATAAGAGAAAGGCTG 6775
Db 7714 GAGCCTCGACTCAGGCTGGCAGAGGGGCTGCTGCTGGCAAAATAAGAGAAAGGCTG 7773
Qy 6776 GAGAGAGCCCTGCTGTGTGAAA 6797
Db 7774 GAGAGAGCCCTGCTGTGTGAAA 7795

RESULT 5
AF328787
LOCUS
DEFINITION
AF328787
ACCESSION
AF328787.1 GI:12656650
VERSION
AF328787.1 GI:12656650
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6704)
Broccardo, C., Osorio, J., Luciani, M.-P., Schriml, L.M., Prades, C.,
Shulenin, S., Arnould, I., Naudin, L., Lafargue, C., Rosier, M.,
Jordan, B., Mattei, M.G., Dean, M., Deneffe, P. and Chimini, G.
Comparative analysis of the promoter structure and genomic
organization of the human and mouse ABCA7 gene encoding a novel
ABCA transporter
Cytogenet. Cell Genet. 92 (3-4), 264-270 (2001)
21328888
11435699
2 (bases 1 to 6704)
Broccardo, C., Osorio, J., Luciani, M.-P., Lynn, S., Prades, C.,
Shulenin, S., Arnould, I., Naudin, L., Lafargue, C., Rosier, M.,
Jordan, B., Mattei, M.G., Dean, M., Deneffe, P. and Chimini, G.
Direct Submission
Submitted (13-DEC-2000) EGC, Aventis-Pharma, 2 Rue Gaston Cremieux,
EVRY, 91 91000, France
FEATURES
Location/Qualifiers
1..6704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1..6704
/gene="ABCA7"
/note="ABCA7 protein"
/codon_start=1
/product="ABC transporter member 7"
/protein_id="AAK00959.1"
/db_xref="GI:12656651"
gene
CDS
117..6557
/gene="ABCA7"
/note="ABCA7 protein"

SGYLLTLVKARLLPTTNEKADTDMEGSDVTRQEKNGSGSRVGTQPLLALVQHVWPG
ARLVEPLHELVLVPTGAHDGSPATLRELDLRLTGVIYSDTSLREIPLKV
VECAADTDMEGSCQOHLCTGLAGLDVTLRLKMPQETALENGEPANETDQSG
PDVGRVQWALTRQQLALLRFLARSRGLFAQIVLPALFVGLAFLVFLIWP
FGHPALRUSPTMYGAQVSPSDAPDGPBARLLLEALOEAGLEBPVVOHSSHRFSA
PEYPAEVAKVLASGNWTPSPSPACQSRPARRLLPDCPAAAGPFPQPAVTGSGEV
VQNLGRNLSDFLVKTYPRLVROGLTKKWNVRYGFSGLGRDPLSPGBLGRSV
EELWALLSPUPGCGALDRVKNLTAWAHSIDAOQSLAKIWNKNGHSMVAFVNRASNAI
LRAHLPPGPBARHAHSITLNLHPLNLTKEQLSEAAALMASSVDVLVSLCVVPANFSVPAS
FTLVLEERVTRAKHLQMGGLSFTLWIGNFWLDMCNLYLPACIIVLIFLAPQQRAY
VAPANLPLLLLLLQWSTPLMYPASFPFSVPSTAYVVLITCINIFIGISMATFV
LEIFSQKQGEVERIILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGRDQPSRLR
WEVGNKLLAWETQGLFLFTLLQHLRQLPQPRVRSPLILGEEDVAVRERWV
OGATGDVLVRLNLTKYRGORMPAVDRLCIGIPBCEGFLGVNGAGKSTLPRMTYG
DTLASGEAVLASHVAREPSAAHLSMGYCPOSDAIPELLTGREHLLEKRLRGVPEA
QVAQTAGSGLVGLWADYPAGTYSGGNCRKLATALLVGDPAVVFLEPTTGMDD
SARRFLWNSLLAVRGRSSVYMLTSHSMECEALCSRLAMVNRFRCLGSPQHLKGRF
AAGHTLTLRVPAARSQPAFAAFVAAEPFGLSREAHGRRLRFLPPLPGRCALARVFGEL
AVHGAHGVDEDFSVSOTMLEEVLFLYFSKQGDDETEQKEAGVGVDPAPGLQHPKRV
SQFLDDPSTAEVTL"

ORIGIN

Query Match 98.3%; Score 6686.4; DB 9; Length 6704;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6693; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy	94	GTGGGATAAGGAATGAGGTTTCAGAAAGGGCGAGGAGTTGCCGCGAGCGCGACG	153
Db	1	GTGGGATAAGGAATGAGGTTTCAGAAAGGGCGAGGAGTTGCCGCGAGCGCGACG	60
Qy	154	TCCTTACGCGCGACGCTGCTGACCTCTCTGCTCCGCTCCCGCCCGCCAGTCTCACATGG	213
Db	61	TCCTTACGCGCGACGCTGCTGACCTCTCTGCTCCGCTCCCGCCCGCCAGTCTCACATGG	120
Qy	214	CTTCTGGACACAGCTGATGCTGCTGCTGGAAGAAATTTTCATGTATCGCGGAGACAGC	273
Db	121	CTTCTGGACACAGCTGATGCTGCTGCTGGAAGAAATTTTCATGTATCGCGGAGACAGC	180
Qy	274	CGGTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	333
Db	181	CGGTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240
Qy	334	TTCCGCACTCCACCCGCTTGAGCACCATGAATGCCACTTCCCAAAACAGCCACTGC	393
Db	241	TTCCGCACTCCACCCGCTTGAGCACCATGAATGCCACTTCCCAAAACAGCCACTGC	300
Qy	394	CATCGCGGCGACCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	453
Db	301	CATCGCGGCGACCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	360
Qy	454	TTCCGCACTCCACCCGCTTGAGCACCATGAATGCCACTTCCCAAAACAGCCACTGC	513
Db	361	TTCCGCACTCCACCCGCTTGAGCACCATGAATGCCACTTCCCAAAACAGCCACTGC	420
Qy	514	TCCTCCGCTGCTAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	573
Db	421	TCCTCCGCTGCTAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Qy	574	TGGTGGCTTAGGGAAGCTGATCGCACGCTGAGGGCTGCACGACGACGCGCCAGCTC	633
Db	481	TGGTGGCTTAGGGAAGCTGATCGCACGCTGAGGGCTGCACGACGACGCGCCAGCTC	540
Qy	634	AACCAACAAAGAGTCTTCCAATGAGCAACCAATGCTGAGATGCTGCGGAGCTGCTGACGT	693
Db	541	AACCAACAAAGAGTCTTCCAATGAGCAACCAATGCTGAGATGCTGCGGAGCTGCTGACGT	600
Qy	694	CACCTGCTGCGCACGGAATCCCTGGGTTGCACTTGGSCCAAGCCACGAGCCCTTGCACA	753
Db	601	CACCTGCTGCGCACGGAATCCCTGGGTTGCACTTGGSCCAAGCCACGAGCCCTTGCACA	660
Qy	754	GCTTGTGGAGGCGCTGAGGACCTGGCCACGAGGCTCTCTGGGCTGCGAGCCTGGTGG	813
Db	661	GCTTGTGGAGGCGCTGAGGACCTGGCCACGAGGCTCTCTGGGCTGCGAGCCTGGTGG	720

Qy	814	AGCTTCGGGCACCTGCTGCAGAGACCCCGAGGACACAGCGGCCCTCTGGAGTTCTGTCTAG	873
Db	721	AGCTTCGGGCACCTGCTGCAGAGACCCCGAGGACACAGCGGCCCTCTGGAGTTCTGTCTAG	780
Qy	874	AGGCCCTCTGCAGTGTTCAGGGGACCTAGCAGACACAGTGGGGCCCTCCCTCAACTGTGTAG	933
Db	781	AGGCCCTCTGCAGTGTTCAGGGGACCTAGCAGACACAGTGGGGCCCTCCCTCAACTGTGTAG	840
Qy	934	AGGCTAGTGACCTGATGGAGCTGGTGGGCGAGGAGCCAGAAATCCGCGCTCTGCCAGACGCA	993
Db	841	AGGCTAGTGACCTGATGGAGCTGGTGGGCGAGGAGCCAGAAATCCGCGCTCTGCCAGACGCA	900
Qy	994	GCCTGAGCCCGGCTCTCTCGGAGCTCATTTGGAGCCCTTGACACACACCCGCTCTCCCGCC	1053
Db	901	GCCTGAGCCCGGCTCTCTCGGAGCTCATTTGGAGCCCTTGACACACACCCGCTCTCCCGCC	960
Qy	1054	TGCTCTGGAGACGCTGAAAGCTCTGATCCTCGGGAAGTACTCTTTTGACACGATACAC	1113
Db	961	TGCTCTGGAGACGCTGAAAGCTCTGATCCTCGGGAAGTACTCTTTTGACACGATACAC	1020
Qy	1114	CTTTTACCGGAGCTCATGGCCAGGTGAAACCGGACCTTTGAGGAGCTCACCCCTGCTGA	1173
Db	1021	CTTTTACCGGAGCTCATGGCCAGGTGAAACCGGACCTTTGAGGAGCTCACCCCTGCTGA	1080
Qy	1174	GGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTTCATCTCATGAACGACA	1233
Db	1081	GGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTTCATCTCATGAACGACA	1140
Qy	1234	GTTCCTCAATGTGGCCATGCTGCAGCGGCTCTGAGATGAGGATGAAGAAAGGAGCAGC	1293
Db	1141	GTTCCTCAATGTGGCCATGCTGCAGCGGCTCTGAGATGAGGATGAAGAAAGGAGCAGC	1200
Qy	1294	CGAGACTGAGGCGCGGACCAATGAGGAGCCCTGCGATCTTCTGGAACCTTGGAGCG	1353
Db	1201	CGAGACTGAGGCGCGGACCAATGAGGAGCCCTGCGATCTTCTGGAACCTTGGAGCG	1260
Qy	1354	GTGGCTACAGCTGGGAGGACGACACGCTGATGTGGGGCACCTGGTGGGCGCCCTGGCC	1413
Db	1261	GTGGCTACAGCTGGGAGGACGACACGCTGATGTGGGGCACCTGGTGGGCGCCCTGGCC	1320
Qy	1414	GAGTGAAGGAGTGCCTGCTCTTGGACAAGCTGGAGCGGACCCCTCAGAGGACCCCTGG	1473
Db	1321	GAGTGAAGGAGTGCCTGCTCTTGGACAAGCTGGAGCGGACCCCTCAGAGGACCCCTGG	1380
Qy	1474	TGTGCGGGCCCTGGAACCTGCTCGGGAAACATGATCTGGGCGGCGCTGCTTCTTGG	1533
Db	1381	TGTGCGGGCCCTGGAACCTGCTCGGGAAACATGATCTGGGCGGCGCTGCTTCTTGG	1440
Qy	1534	GACCTGAGGACTCTTTCAGACCCACAGACACCCCAACCCAGACCTGGGCGCGGCGCACG	1593
Db	1441	GACCTGAGGACTCTTTCAGACCCACAGACACCCCAACCCAGACCTGGGCGCGGCGCACG	1500
Qy	1594	TGCGCATCAAAATCCGCATGGACATTTGAGTGTGTCAAGAGGACCAATAAGATCAGGACA	1653
Db	1501	TGCGCATCAAAATCCGCATGGACATTTGAGTGTGTCAAGAGGACCAATAAGATCAGGACA	1560
Qy	1654	GGTTTGGGACCCCTGGCCAGCGCGGACCCCTGACCCGACCTGCGGTAGCTGTGGGGCG	1713
Db	1561	GGTTTGGGACCCCTGGCCAGCGCGGACCCCTGACCCGACCTGCGGTAGCTGTGGGGCG	1620
Qy	1714	GCTTCGTGTACCTGCAAGACCTGGTGGAGCTGACGCCCTGCGCGTCTCAGCGGCGCCA	1773
Db	1621	GCTTCGTGTACCTGCAAGACCTGGTGGAGCTGACGCCCTGCGCGTCTCAGCGGCGCCA	1680
Qy	1774	ACCCCGGGCGCGGCTCTTACCTGACAGAGATGCCCTTATCCGTGCTATGTGGACAGCTGT	1833
Db	1681	ACCCCGGGCGCGGCTCTTACCTGACAGAGATGCCCTTATCCGTGCTATGTGGACAGCTGT	1740
Qy	1834	TCTGTGTGTGTGACCGGCTGCTGCCGCTCTTCTGTGAGCTGGCTGCTGATCTCTCCG	1893
Db	1741	TCTGTGTGTGTGACCGGCTGCTGCCGCTCTTCTGTGAGCTGGCTGCTGATCTCTCCG	1800

QY 1894 TGACACTGACAGTGAAGCCGCTGTGTCGCGGAGAAAGAGACGCGGCTGCGGGACACATGTC 1953
DB 1801 TGACACTGACAGTGAAGCCGCTGTGTCGCGGAGAAAGAGACGCGGCTGCGGGACACATGTC 1860
QY 1954 GCGGCATGGGGCTCAGCCGCGGCTGTCTGCTAGGCTAGGCTGTCTCAGCTGCTCGGGC 2013
DB 1861 GGGCCATGGGGCTCAGCCGCGGCTGTCTGCTAGGCTAGGCTGTCTCAGCTGCTCGGGC 1920
QY 2014 CTTCTCTGCTCAGCCGCGGCTGTGTTCTGGTCTCAAGCTGGGGACATCTCCCT 2073
DB 1921 CTTCTCTGCTCAGCCGCGGCTGTGTTCTGGTCTCAAGCTGGGGACATCTCCCT 1980
QY 2074 ACAGCCACCCGGGCTGTGTTCTGCTTGGCAGCCTTTCGCGTGGCCAGCGTGACCC 2133
DB 1981 ACAGCCACCCGGGCTGTGTTCTGCTTGGCAGCCTTTCGCGTGGCCAGCGTGACCC 2040
QY 2134 AGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAACTGGCTGGGCTGCGGGGCC 2193
DB 2041 AGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAACTGGCTGGGCTGCGGGGCC 2100
QY 2194 TGGCCTACTTCTCCCTCTACCTGCCCCTAGCTGCTGTGTGTGGCTTGGCGGACCCGCTGTC 2253
DB 2101 TGGCCTACTTCTCCCTCTACCTGCCCCTAGCTGCTGTGTGTGGCTTGGCGGACCCGCTGTC 2160
QY 2254 CCGGGGTGGCGGCTGCGCGGAGCCTGCTGTGTCGCCGCTGGCCTTTCGCGTTCGCTGCG 2313
DB 2161 CCGGGGTGGCGGCTGCGCGGAGCCTGCTGTGTCGCCGCTGGCCTTTCGCGTTCGCGTTCG 2220
QY 2314 AGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAACTGGCTGGGCTGCGGGGCC 2373
DB 2221 AGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAACTGGCTGGGCTGCGGGGCC 2280
QY 2374 GGCCTACCGGACAGCTTCTTACGCTGGCCGAGGCTCTTGGCTTCTGCTGTGACGCGG 2433
DB 2281 GGCCTACCGGACAGCTTCTTACGCTGGCCGAGGCTCTTGGCTTCTGCTGTGACGCGG 2340
QY 2434 CGCTCTACGGCTTCGCCACCTGGTACCTTGAAGCTGTGTGCCAGGCGCAGTACGGGATCC 2493
DB 2341 CGCTCTACGGCTTCGCCACCTGGTACCTTGAAGCTGTGTGCCAGGCGCAGTACGGGATCC 2400
QY 2494 CTGACCACTGGAATTTTCTTTTCGGAGGAGTACTGTGTGGGAGCTTGGGCTCCCAAGA 2553
DB 2401 CTGAACCACTGGAATTTTCTTTTCGGAGGAGTACTGTGTGGGAGCTTGGGCTCCCAAGA 2460
QY 2554 GTCCAGGCCCTTGGCCCAACCCGCTGGACCCAAAGGTGCTGTGAAGAGGACCGGCCG 2613
DB 2461 GTCCAGGCCCTTGGCCCAACCCGCTGGACCCAAAGGTGCTGTGAAGAGGACCGGCCG 2520
QY 2614 GCCTGAGTCTTGGCGTATCCGTTTCGAGCCTTGGAGAGCGCTTTCCTGGAAGCCGCGAGC 2673
DB 2521 GCCTGAGTCTTGGCGTCTCCGTTTCGAGCCTTGGAGAGCGCTTTCCTGGAAGCCGCGAGC 2580
QY 2674 CAGCCCTCGGGGGCTCAGCTTGAATTTCTACAGGGCCACATCACCGCTTCTTGGGCC 2733
DB 2581 CAGCCCTCGGGGGCTCAGCTTGAATTTCTACAGGGCCACATCACCGCTTCTTGGGCC 2640
QY 2734 ACAACGGGGCGGCAAGACCAACCTGCTCATCTTGAAGTGGCTTTCCTCCACCCAGTG 2793
DB 2641 ACAACGGGGCGGCAAGACCAACCTGCTCATCTTGAAGTGGCTTTCCTCCACCCAGTG 2700
QY 2794 GTGGCTTCTGCTTATCTTGGGCCACGACGCTCGCTCCAGCATGGCCGCTTTCGCGGCC 2853
DB 2701 GTGGCTTCTGCTTATCTTGGGCCACGACGCTCGCTCCAGCATGGCCGCTTTCGCGGCC 2760
QY 2854 ACCTGGGCGTCTGCTCAGTACACAGTGTGTTTGAATGCTGACCGTGAAGGACG 2913
DB 2761 ACCTGGGCGTCTGCTCAGTACACAGTGTGTTTGAATGCTGACCGTGAAGGACG 2820
QY 2914 TCTGCTTCTATGGGGCTGGAAGGCTGTGAGTGGCGCTGTAGTGGGCCCCCAGAGGACC 2973
DB 2821 TCTGCTTCTATGGGGCTGGAAGGCTGTGAGTGGCGCTGTAGTGGGCCCCCAGAGGACC 2880
QY 2974 GTCTGCTGACGAGTGTGGGGCTGTGTTCTCCAAAGCAGAGTGTGACACTCGCCACCTCTCTG 3033

DB 2881 GTCTGCTGACGAGTGTGGGGCTGTGTTCTCCAAAGCAGAGTGTGACAGTGTGCGCACCTCTCTG 2940
QY 3034 GTGGGATCAACGGAAGCTGTCCGTGGCCATTTGCTTGTGGCGGCTCCCAAGTTGTTA 3093
DB 2941 GTGGGATCAACGGAAGCTGTCCGTGGCCATTTGCTTGTGGCGGCTCCCAAGTTGTTA 3000
QY 3094 TCCTGGACGAGCTACCGCTGGCGCTGATCTCTGCTTCCCGCGCGGTATTTGGGAGCTGC 3153
DB 3001 TCCTGGACGAGCTACCGCTGGCGCTGATCTCTGCTTCCCGCGCGGTATTTGGGAGCTGC 3060
QY 3154 TGCTCAAAATACGGAAGGTGCAAGCTGATCTCTTCCACCCACCACTGAGTGAAGGACG 3213
DB 3061 TGCTCAAAATACGGAAGGTGCAAGCTGATCTCTTCCACCCACCACTGAGTGAAGGACG 3120
QY 3214 AGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTCCC 3273
DB 3121 AGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTCCC 3180
QY 3274 CACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCGCGC 3333
DB 3181 CACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCGCGC 3240
QY 3334 TGCCCTTGACCAACCAATGAGAAAGGTGACACTGACATGGAGGCGAGTGTGACACACAGGC 3393
DB 3241 TGCCCTTGACCAACCAATGAGAAAGGTGACACTGACATGGAGGCGAGTGTGACACACAGGC 3300
QY 3394 AGGAAAGAAATGGCAGCCAGGCGCAGAGTGGCACCTCCAGCTGCTGGCCCTGG 3453
DB 3301 AGGAAAGAAATGGCAGCCAGGCGCAGAGTGGCACCTCCAGCTGCTGGCCCTGG 3360
QY 3454 TACAGCACTGGTGTCCCGGGCAGCGGTGGTGGAGGAGCTGCCACACAGAGCTGGTGTGG 3513
DB 3361 TACAGCACTGGTGTCCCGGGCAGCGGTGGTGGAGGAGCTGCCACACAGAGCTGGTGTGG 3420
QY 3514 TGCTGCCCTACACGGGTGCCCATGACGCGAGCTTGCGCCACACTCTTTCGAGAGCTAGACA 3573
DB 3421 TGCTGCCCTACACGGGTGCCCATGACGCGAGCTTGCGCCACACTCTTTCGAGAGCTAGACA 3480
QY 3574 CGCGGCTGGCGAGCTGAGGCTCACTGGTACGGGATCTCCGACACACAGCTTCGAGGAGA 3633
DB 3481 CGCGGCTGGCGAGCTGAGGCTCACTGGTACGGGATCTCCGACACACAGCTTCGAGGAGA 3540
QY 3634 TCTTCTGAAAGTGGTGGAGGAGTGTCTGCGGACACAGATATGGAGATGGCAGCTGCG 3693
DB 3541 TCTTCTGAAAGTGGTGGAGGAGTGTCTGCGGACACAGATATGGAGATGGCAGCTGCG 3600
QY 3694 GGCAGCACTATGACAGGCAATGCTGGCTTAGACGTAAACCTGCGGCTCAAGATGCCGC 3753
DB 3601 GGCAGCACTATGACAGGCAATGCTGGCTTAGACGTAAACCTGCGGCTCAAGATGCCGC 3660
QY 3754 CACAGGAGACGCTGGAGAACGGGNAACAGCTGGGTGAGCCCGCAGAGCTGACGAG 3813
DB 3661 CACAGGAGACGCTGGAGAACGGGNAACAGCTGGGTGAGCCCGCAGAGCTGACGAG 3720
QY 3814 GCTCTGGGCCACAGCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGCAGAGCTCC 3873
DB 3721 GCTCTGGGCCACAGCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGCAGAGCTCC 3780
QY 3874 AGGCCCTGCTTCTCAAAGCGCTTCTGCTTGGCCCGCGCAGCGCGCGCTGTTGCGCC 3933
DB 3781 AGGCCCTGCTTCTCAAAGCGCTTCTGCTTGGCCCGCGCAGCGCGCGCTGTTGCGCC 3840
QY 3934 AGATGCTGCTGCTTCTTCTGCGGCTGGGCTCGTGTTCAGCTCATCGTGCCTC 3993
DB 3841 AGATGCTGCTGCTTCTTCTGCGGCTGGGCTCGTGTTCAGCTCATCGTGCCTC 3900
QY 3994 CTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGATGACGCTCAGGCTGCTCCT 4053
DB 3901 CTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGATGACGCTCAGGCTGCTCCT 3960
QY 4054 TCTTCAGTGAAGGACCGCCCGAGGGACCTGGAAGTGGCCCGCTGCTGAGGCGCTGCTGC 4113

Db 361 ACGTGCTGCTAGGGAAGCTGATGCCACGCTGAGGGCTGCACGACGACGGCCAG 420
Qy 630 CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCCATGCTGGATGTCGGGAGCTGCTG 689
Db 421 CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCCATGCTGGATGTCGGGAGCTGCTG 480
Qy 690 ACGTCACTGCTGCGCACCGGAATCCCTGGGGTGGCACTGGGCCAAAGCCACAGAGCCCTTG 749
Db 481 ACGTCACTGCTGCGCACCGGAATCCCTGGGGTGGCACTGGGCCAAAGCCACAGAGCCCTTG 540
Qy 750 CACAGCTTGTGGAGGCGCTGAGAGACCTGGCCAGAGATCTCTGGCGCTGGCGAGCCTG 809
Db 541 CACAGCTTGTGGAGGCGCTGAGAGACCTGGCCAGAGATCTCTGGCGCTGGCGAGCCTG 600
Qy 810 GTGGAGCTTCGGGACATGCTGCAGAGACCCGAGGGACACGCGCCCTGGAGTTGCTG 869
Db 601 GTGGAGCTTCGGGACATGCTGCAGAGACCCGAGGGACACGCGCCCTGGAGTTGCTG 660
Qy 870 TCAGAGCCCTCTGCACTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGG 929
Db 661 TCAGAGCCCTCTGCACTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGG 720
Qy 930 TACAGGCTAGTGAATGAGTGTGGAGCTGGTGGGAGAGCCAGAAATCCGCCCTGCCAGAC 989
Db 721 TACAGGCTAGTGAATGAGTGTGGAGCTGGTGGGAGAGCCAGAAATCCGCCCTGCCAGAC 780
Qy 990 AGCAGCTGAGCCCGCTGCTGGAGCTGATTTGGAGCCCTTGACAGCCACCGCTGTCC 1049
Db 781 AGCAGCTGAGCCCGCTGCTGGAGCTGATTTGGAGCCCTTGACAGCCACCGCTGTCC 840
Qy 1050 CGCTGCTCTGGAGACCCCTGAAGCCTCTGATCTCGGGAAGCTTACTTTTGACACAGAT 1109
Db 841 CGCTGCTCTGGAGACCCCTGAAGCCTCTGATCTCGGGAAGCTTACTTTTGACACAGAT 900
Qy 1110 ACACCTTTTACCGGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAGCTCACCCTG 1169
Db 901 ACACCTTTTACCGGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAGCTCACCCTG 960
Qy 1170 CTGAGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACCTTCATGAAC 1229
Db 961 CTGAGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACCTTCATGAAC 1020
Qy 1230 GACAGTTTCAATGTGGCCATGCTGCAGCGCTCTGTCAGATGACAGGATGAAGAAGG 1289
Db 1021 GACAGTTTCAATGTGGCCATGCTGCAGCGCTCTGTCAGATGACAGGATGAAGAAGG 1080
Qy 1290 CAGCCAGACCTGAGGCGCGGACCAACATGAGGSCCTGCGATCTTTTGGACCCCTGG 1349
Db 1081 CAGCCAGACCTGAGGCGCGGACCAACATGAGGSCCTGCGATCTTTTGGACCCCTGG 1140
Qy 1350 AGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGACCCCTGGTGGGACGCTG 1409
Db 1141 AGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGACCCCTGGTGGGACGCTG 1200
Qy 1410 GGCGAGTGAAGGAGTCCCTGCTCGACAACTGGAGCGGACCCCTCAGAGGACGCC 1469
Db 1201 GGCGAGTGAAGGAGTCCCTGCTCGACAACTGGAGCGGACCCCTCAGAGGACGCC 1260
Qy 1470 CTGCTGTGGGGGCCCTGCAACTGCTCGCGGAAACATCGATTTCTGGGGCGGCGCTGCTTC 1529
Db 1261 CTGCTGTGGGGGCCCTGCAACTGCTCGCGGAAACATCGATTTCTGGGGCGGCGCTGCTTC 1320
Qy 1530 TTGGGACCTGAGGACTCTTCAGACCCCAAGAGCACCCCAACCCAGACCTTGGGCCCGCGC 1589
Db 1321 TTGGGACCTGAGGACTCTTCAGACCCCAAGAGCACCCCAACCCAGACCTTGGGCCCGCGC 1380
Qy 1590 CAGTGGCATCAAAATCCGATGACATTTGAGTGTACAGGACCAATAGATCAGG 1649
Db 1381 CAGTGGCATCAAAATCCGATGACATTTGAGTGTACAGGACCAATAGATCAGG 1440
Qy 1650 GACAGGTTTTGGACCCCTGGCCAGCGGACCCCTGACCGACCTTCGCTACGCTGG 1709
Db 1441 GACAGGTTTTGGACCCCTGGCCAGCGGACCCCTGACCGACCTTCGCTACGCTGG 1500

Qy 1710 GGCGGCTTCTGTATCCTGCAAGACCTGGTGGAGCGTGCAGCCGTCGCGTGTCTCAGCGC 1769
Db 1501 GGCGGCTTCTGTATCCTGCAAGACCTGGTGGAGCGTGCAGCCGTCGCGTGTCTCAGCGC 1560
Qy 1770 GCCAACCCCGGGCCGGCTCTACCTGCAGCAGATCCCTATCCGTGTATGTGGACGAC 1829
Db 1561 GCCAACCCCGGGCCGGCTCTACCTGCAGCAGATCCCTATCCGTGTATGTGGACGAC 1620
Qy 1830 GTGTTTCTGGGTGTCTGAGCGCGTGCCTGCGCTCTTCTGACGCTGGCCTGGATCTAC 1889
Db 1621 GTGTTTCTGGGTGTCTGAGCGCGTGCCTGCGCTCTTCTGACGCTGGCCTGGATCTAC 1680
Qy 1890 TCCGTGACACTGACAGTGAAGCCGTGTGCGGAGAGGAGACGCGGCTGCGGACACC 1949
Db 1681 TCCGTGACACTGACAGTGAAGCCGTGTGCGGAGAGGAGACGCGGCTGCGGACACC 1740
Qy 1950 ATCGCGCCATGGGGCTCAGCCGCGGCTGCTCTGCTAGGCTGGTTCCTCAGCTGCTC 2009
Db 1741 ATCGCGCCATGGGGCTCAGCCGCGGCTGCTCTGCTAGGCTGGTTCCTCAGCTGCTC 1800
Qy 2010 GGCCCTTCTGCTCAGCGCGGCTGCTGCTCTGCTGCTCAAGCTGGGGACATCCTC 2069
Db 1801 GGCCCTTCTGCTCAGCGCGCCTGCTGCTCTGCTGCTCAAGCTGGGGACATCCTC 1860
Qy 2070 CCTCAGACACCCCGGCGGTGCTCTGCTGCTGCGAGCCTTCGCGGTGGCCACGCTG 2129
Db 1861 CCTCAGACACCCCGGCGGTGCTCTGCTGCTGCGAGCCTTCGCGGTGGCCACGCTG 1920
Qy 2130 ACCAGAGCTTCTGCTCAGCGCGCTTCTCTCCCGGCCAACCTGGCTGCGGCTGCGGC 2189
Db 1921 ACCAGAGCTTCTGCTCAGCGCGCTTCTCTCCCGGCCAACCTGGCTGCGGCTGCGGC 1980
Qy 2190 GGCTGGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2249
Db 1981 GGCTGGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Qy 2250 CTGCCCGGGTGGCGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2309
Db 2041 CTGCCCGGGTGGCGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2310 TGGAGAGCCTGCTGCTGCTGAGGAGCAGGGGCGAGGCGCAGTGGCAACAAGTGGGC 2369
Db 2101 TGGAGAGCCTGCTGCTGCTGAGGAGCAGGGGCGAGGCGCAGTGGCAACAAGTGGGC 2160
Qy 2370 ACCCGGCTTACGGCAGAGCTTTCAGCTGCGGCCAGGCTCTGCTGCTGCTGCTGCTG 2429
Db 2161 ACCCGGCTTACGGCAGAGCTTTCAGCTGCGGCCAGGCTCTGCTGCTGCTGCTGCTG 2220
Qy 2430 GGCGGCTCTTACGGCTCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
Db 2221 GGCGGCTCTTACGGCTCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
Qy 2490 ATCCCTGAACCAATGGAATTTTCTTTTTCGAGAGAGCTTCTGCTGCTGCTGCTGCTG 2549
Db 2281 ATCCCTGAACCAATGGAATTTTCTTTTTCGAGAGAGCTTCTGCTGCTGCTGCTGCTG 2340
Qy 2550 AAGAGTCCAGCCCTTTCGCCCCACCCCGCTGGAACCAAGGCTGCTGCTGCTGCTGCTG 2609
Db 2341 AAGAGTCCAGCCCTTTCGCCCCACCCCGCTGGAACCAAGGCTGCTGCTGCTGCTGCTG 2400
Qy 2610 CCGGGCTGAGTCTGCGGCTATCCGTTGCGAGCTGGAAGCGCTTCTCTGGAAGCCCG 2669
Db 2401 CCGGGCTGAGTCTGCGGCTATCCGTTGCGAGCTGGAAGCGCTTCTCTGGAAGCCCG 2460
Qy 2670 CAGCAGCCCTGCGGGGCTCAGCTGGACTTCTACAGGGGCCACATCAGCGCTTCTCTG 2729
Db 2461 CAGCAGCCCTGCGGGGCTCAGCTGGACTTCTACAGGGGCCACATCAGCGCTTCTCTG 2520
Qy 2730 GGCCACAAAGGGCGCGGCAAGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2789
Db 2521 GGCCACAAAGGGCGCGGCAAGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580

Qy	2790	AGTGGTGGCTCTGCCTTCATCTGGGCGACGACGTCCGCTCAGCATGGCGCGCATCCGG	2849
Db	2581	AGTGGTGGCTCTGCCTTCATCTGGGCGACGACGTCCGCTCAGCATGGCGCGCATCCGG	2640
Qy	2850	CCCCACCTGGGGCTCTGCTCCTCAGTACAAAGTCTGTTTGACATCTGACCGTGGACGAG	2909
Db	2641	CCCCACCTGGGGCTCTGCTCCTCAGTACAAAGTCTGTTTGACATCTGACCGTGGACGAG	2700
Qy	2910	CACGCTCTGTTCTATGGGCGGCTGAAGGCTCTGAGTGCCTGTAGTGGGCGGCCGAGCAG	2969
Db	2701	CACGCTCTGTTCTATGGGCGGCTGAAGGCTCTGAGTGCCTGTAGTGGGCGGCCGAGCAG	2760
Qy	2970	GACCGTCTGCTGACAGATGGGGCTGTCTTCCAAAGCAGAGTGTGCACATCGCCACCTC	3029
Db	2761	GACCGTCTGCTGACAGATGGGGCTGTCTTCCAAAGCAGAGTGTGCACATCGCCACCTC	2820
Qy	3030	TCTGTGGATGCAACGGAAGCTGTCCGTGGCCATTGGCTTTGTGGGCGGCTCCCAAGTT	3089
Db	2821	TCTGTGGATGCAACGGAAGCTGTCCGTGGCCATTGGCTTTGTGGGCGGCTCCCAAGTT	2880
Qy	3090	GTTATCTCTGGACGACCTACGCTGGCGGTGGATCCTGCTTCCCGCGCGGTATTTGGGAG	3149
Db	2881	GTTATCTCTGGACGACCTACGCTGGCGGTGGATCCTGCTTCCCGCGCGGTATTTGGGAG	2940
Qy	3150	CTGCTGCTCAAAATACCGAAGAGTTCGACGCTGATCTCTCCACCCACCACTGGATGAG	3209
Db	2941	CTGCTGCTCAAAATACCGAAGAGTTCGACGCTGATCTCTCCACCCACCACTGGATGAG	3000
Qy	3210	GCAGAGCTGCTGGGAGCCGTGTGCTGTGGCTGTGGGAGTGGCGGCTTGTGCTGTGGC	3269
Db	3001	GCAGAGCTGCTGGGAGCCGTGTGCTGTGGCTGTGGGAGTGGCGGCTTGTGCTGTGGC	3060
Qy	3270	TCCCCACTCTTCTCGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC	3329
Db	3061	TCCCCACTCTTCTCGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC	3120
Qy	3310	CGCTGCCCCTGACCAACAATAGAAAGCTGACACTGACATGAGAGGCAAGTGTGGACACC	3389
Db	3121	CGCTGCCCCTGACCAACAATAGAAAGCTGACACTGACATGAGAGGCAAGTGTGGACACC	3180
Qy	3390	AGGCAGGAAAGAAATGGGAGCCAGGCGACGAGTCCGACCTCTCAGCTCTCGGCC	3449
Db	3181	AGGCAGGAAAGAAATGGGAGCCAGGCGACGAGTCCGACCTCTCAGCTCTCGGCC	3240
Qy	3450	CTGGTACAGCACTGGGTGCCCGGGGCAAGGCTGTGGAGGAGCTGCCACAGAGCTGTG	3509
Db	3241	CTGGTACAGCACTGGGTGCCCGGGGCAAGGCTGTGGAGGAGCTGCCACAGAGCTGTG	3300
Qy	3510	CTGGTGTGCCCCATACACGGGTGCCATGACGGCAGCTTCCGCCACACTCTTCCGAGAGCTA	3569
Db	3301	CTGGTGTGCCCCATACACGGGTGCCATGACGGCAGCTTCCGCCACACTCTTCCGAGAGCTA	3360
Qy	3570	GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAG	3629
Db	3361	GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAG	3420
Qy	3630	GAGATCTTCTCAAGGCTGGAGAGTGTGTGCGGACACAGATATGAGAGATGGCAGC	3689
Db	3421	GAGATCTTCTCAAGGCTGGAGAGTGTGTGCGGACACAGATATGAGAGATGGCAGC	3480
Qy	3690	TGCGGGCAGCACCTATGACAGGCAATGCTGGCCCTAGACGTAAACCTCGGCTCAAGATG	3749
Db	3481	TGCGGGCAGCACCTATGACAGGCAATGCTGGCCCTAGACGTAAACCTCGGCTCAAGATG	3540
Qy	3750	CCGCACAGGAGACAGCCCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCAGAGCTGAC	3809
Db	3541	CCGCACAGGAGACAGCCCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCAGAGCTGAC	3600
Qy	3810	CAGGCTCTGGGCCAGAGCCGCTGGGCGGGTACAGAGGCTGGGCACTGACCCCGCAGCAG	3869
Db	3601	CAGGCTCTGGGCCAGAGCCGCTGGGCGGGTACAGAGGCTGGGCACTGACCCCGCAGCAG	3660
Qy	3870	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCAGCGCGCGCGCTGTTC	3929

Db	3651	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGCCTCCGCGAGCCGCGCGCTGTTC	3720
Qy	3930	GCCAGATCTGCTGCTGCCCTCTCTTCTGGGCGCTTGGCCCTCGTGTTCAGCCCTCATCTG	3989
Db	3721	GCCAGATCTGCTGCTGCCCTCTCTTCTGGGCGCTTGGCCCTCGTGTTCAGCCCTCATCTG	3780
Qy	3990	CCTCTTCTGGGCACTACCCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTG	4049
Db	3781	CCTCTTCTGGGCACTACCCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTG	3840
Qy	4050	TCCTTCTTCACTGAGGAGCCGCCAGGGACCCCTGGAAGTGCCTGGCTCTCGAGGCGCTG	4109
Db	3841	TCCTTCTTCACTGAGGAGCCGCCAGGGACCCCTGGAAGTGCCTGGCTCTCGAGGCGCTG	3900
Qy	4110	CTGCAGGAGGACAGCTTGGAGGAGCCCTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGT	4169
Db	3901	CTGCAGGAGGACAGCTTGGAGGAGCCCTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGT	3960
Qy	4170	CCAGAAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGACCCCAAGTCT	4229
Db	3961	CCAGAAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGACCCCAAGTCT	4020
Qy	4230	CCATCCCGAGCTTGCAGTGTAGCCAGCCCGGTGCGCGCGCTGCTGCCGAGTGCCTG	4289
Db	4021	CCATCCCGAGCTTGCAGTGTAGCCAGCCCGGTGCGCGCGCTGCTGCCGAGTGCCTG	4080
Qy	4290	GCTCAGCTGTGTCTCCCTCCGCGCAGGAGTGCAGGCTCTTGGGGAGTGGTTCAG	4349
Db	4081	GCTCAGCTGTGTCTCCCTCCGCGCAGGAGTGCAGGCTCTTGGGGAGTGGTTCAG	4140
Qy	4350	AACCTGACAGCCCGGAACCTGCTGACTTCTGGTCAAGACTACCCCGGCTTGGTGGC	4409
Db	4141	AACCTGACAGCCCGGAACCTGCTGACTTCTGGTCAAGACTACCCCGGCTTGGTGGC	4200
Qy	4410	CAGGCTTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACGAGGCTTCTCGTGGG	4469
Db	4201	CAGGCTTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACGAGGCTTCTCGTGGG	4260
Qy	4470	GGCGAGACCCAGGCTTCCCTCGGCGCAAGTGGCGCTCAGTGGAGGAGTGTGG	4529
Db	4261	GGCGAGACCCAGGCTTCCCTCGGCGCAAGTGGCGCTCAGTGGAGGAGTGTGG	4320
Qy	4530	GCCTGCTGAGTCCCTCGCTGGCGGGCCCTCGACCGTGTCTTGAAGAACTTCACAGCC	4589
Db	4321	GCCTGCTGAGTCCCTCGCTGGCGGGCCCTCGACCGTGTCTTGAAGAACTTCACAGCC	4380
Qy	4590	TGGGCTCACAGCTTGAAGTCTCAGGACAGTCTCAAGATCTGGTTCACAAAGGCTGG	4649
Db	4381	TGGGCTCACAGCTTGAAGTCTCAGGACAGTCTCAAGATCTGGTTCACAAAGGCTGG	4440
Qy	4650	CACCTCCATGTGGCTTGTTCACCGAGCCAGCAAGCAATCCTCCGTGCTCACCTGCC	4709
Db	4441	CACCTCCATGTGGCTTGTTCACCGAGCCAGCAAGCAATCCTCCGTGCTCACCTGCC	4500
Qy	4710	CCAGGCGCGGCCCGCCACAGCATCACACACTCAACACCCCTTGAACCTCAC	4769
Db	4501	CCAGGCGCGGCCCGCCACAGCATCACACACTCAACACCCCTTGAACCTCAC	4560
Qy	4770	AAGGAGCAGTGTCTGAGGCTGCACTGATGGCTTCTCGGTGGAGCTCTCGTCTCCATC	4829
Db	4561	AAGGAGCAGTGTCTGAGGCTGCACTGATGGCTTCTCGGTGGAGCTCTCGTCTCCATC	4620
Qy	4830	TGCTGTGTCTTGTGCTTGTCCCGGCGAGCTTCACTTGTCTCTCATTTGAGGAG	4889
Db	4621	TGCTGTGTCTTGTGCTTGTCCCGGCGAGCTTCACTTGTCTCTCATTTGAGGAG	4680
Qy	4890	CGAGTCAACCCGAGCAAGCAGCTCAGCTCATTGGGGGCTTGTCCCCCAGCCTCTACTGG	4949
Db	4681	CGAGTCAACCCGAGCAAGCAGCTCAGCTCATTGGGGGCTTGTCCCCCAGCCTCTACTGG	4740
Qy	4950	CTTGGCAACTTCTCTGGGACATGTAACTTATTTGGTGGCAGCATGATCGTGTGTCTC	5009

Query Match		95.2%;	Score 6478.4;	DB 6;	Length 6522;
Best Local Similarity		99.7%;	Pred. No. 0;		
Matches 6501; Conservative		0;	Mismatches 17;	Indels 3;	Gaps 1;
Qy	210	ATGGCCCTCTTGACACAGCTGATGCTGCTCTGGAGAAATTTTCATGATATCGCCGAGGA	269		
Db	1	ATGGCCCTCTTGACACAGCTGATGCTGCTCTGGAGAAATTTTCATGATATCGCCGAGGA	60		
Qy	270	CAGCCGGTCCAGCTCTCTGATCGAAATTTGCTGTGGCCCTCTCTTCTCTTTCATCTGCTG	329		
Db	61	CAGCCGGTCCAGCTCTCTGATCGAAATTTGCTGTGGCCCTCTCTTCTCTTTCATCTGCTG	120		
Qy	330	GCTGTTGCGCACTCCACCCGCCCTTGAGAGCAACATGAATGCCACTTCCCAACAGCCA	389		
Db	121	GCTGTTGCGCACTCCACCCGCCCTTGAGAGCAACATGAATGCCACTTCCCAACAGCCA	180		
Qy	390	CTGCCATCGGGGGGACCGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	449		
Db	181	CTGCCATCGGGGGGACCGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240		
Qy	450	TGCTTTCCGCACTGACACCGGGCGAGGAGCCCGGGCGCTGAGCAACTTCAACGACTCC	509		
Db	241	TGCTTTCCGCACTGACACCGGGCGAGGAGCCCGGGCGCTGAGCAACTTCAACGACTCC	300		
Qy	510	CTGCTCTCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGAGGGGCGAGTCCCAAGG	569		
Db	301	CTGCTCTCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGAGGGGCGAGTCCCAAGG	360		
Qy	570	ACGCTGGCTGGCTAGGGAAGCTGATGCCAGCTGAGGCTGACGAGCAACGCGCCAG	629		
Db	361	ACGCTGGCTGGCTAGGGAAGCTGATGCCAGCTGAGGCTGACGAGCAACGCGCCAG	420		
Qy	630	CCTCAACCAACCAACAGCTCTCCACTGGAAACCAACCCATGCTGGATGTCGGAGCTGCTG	689		
Db	421	CCTCAACCAACCAACAGCTCTCCACTGGAAACCAACCCATGCTGGATGTCGGAGCTGCTG	480		
Qy	690	ACGTCACTGCTGCGCAAGAAATCTCTGGGGTTGGCACTGGGGCAAGCCCAAGAGCCCTTG	749		
Db	481	ACGTCACTGCTGCGCAAGAAATCTCTGGGGTTGGCACTGGGGCAAGCCCAAGAGCCCTTG	540		
Qy	750	CACAGCTTTGAGGGCGCTGAGGACCTGGCCAGAGGCTCCTGGCGCTGGCGAGCTG	809		
Db	541	CACAGCTTTGAGGGCGCTGAGGACCTGGCCAGAGGCTCCTGGCGCTGGCGAGCTG	600		
Qy	810	GTGAGCTTTGGGCACTGCTGCAGAGACCCCGAGGGAACAGCGGCCCTCTGGAGTTGCTG	869		
Db	601	GTGAGCTTTGGGCACTGCTGCAGAGACCCCGAGGGAACAGCGGCCCTCTGGAGTTGCTG	660		
Qy	870	TCAGAGGCCCTCTGCAGTGTGAGGGGAACCTAGCAGCAAGTGGGGCCCTCCCTCAACTGG	929		
Db	661	TCAGAGGCCCTCTGCAGTGTGAGGGGAACCTAGCAGCAAGTGGGGCCCTCCCTCAACTGG	720		
Qy	930	TACGAGGCTAGTACCTGATGAGCTGTGGGCAAGGAGCCAGAACTCCGCCCTGCGAGAC	989		
Db	721	TACGAGGCTAGTACCTGATGAGCTGTGGGCAAGGAGCCAGAACTCCGCCCTGCGAGAC	780		
Qy	990	AGCAGCTGAGCCCGCTGCTGAGCTGATTTGAGCCCTTGGACAGCAACCCCGTGTCC	1049		
Db	781	AGCAGCTGAGCCCGCTGCTGAGCTGATTTGAGCCCTTGGACAGCAACCCCGTGTCC	840		
Qy	1050	CGCTGCTCTGGAGACCGCTGAAGCTCTGATCTCTGGGAAGCTACTCTTTGGACCAAGAT	1109		
Db	841	CGCTGCTCTGGAGACCGCTGAAGCTCTGATCTCTGGGAAGCTACTCTTTGCAACAGAT	900		
Qy	1110	ACACCTTTTACCGGAAGCTCATGGCCAGAGTGAACCGGACCTTCGAGGAGCTCACCTG	1169		
Db	901	ACACCTTTTACCGGAAGCTCATGGCCAGAGTGAACCGGACCTTCGAGGAGCTCACCTG	960		
Qy	1170	CTGAGGAGTGTCCGGGAGGTGTGGAGATGCTGGAGCCCGGATCTTCACTTTCATGAAC	1229		
Db	961	CTGAGGAGTGTCCGGGAGGTGTGGAGATGCTGGAGCCCGGATCTTCACTTTCATGAAC	1020		
Qy	1230	GACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCCTGCAAGATGACAGGATGAAGGAAGG	1289		

Db	1021	GACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCCTGCAAGATGACAGGATGAAGGAAGG	1080		
Qy	1290	CAGCCCAACAGCTTGGAGGGCGGACACATGAGAGGCCCTGCGATCTTCTTCTGACCTGGG	1349		
Db	1081	CAGCCCAACAGCTTGGAGGGCGGACACATGAGAGGCCCTGCGATCTTCTTCTGACCTGGG	1140		
Qy	1350	AGCGTGGCTACAGCTGGCAGGACGACACATGATGTGGGGCACTGTGTGGGCAACGCTG	1409		
Db	1141	AGCGTGGCTACAGCTGGCAGGACGACACATGATGTGGGGCACTGTGTGGGCAACGCTG	1200		
Qy	1410	GCCCGAGTGAAGAGTGTCTGTCTTGGACAAGCTGGAGGGGCAACCTCTAGAGGAGGCC	1469		
Db	1201	GCCCGAGTGAAGAGTGTCTGTCTTGGACAAGCTGGAGGGGCAACCTCTAGAGGAGGCC	1260		
Qy	1470	CTGCTGTGCGGGGCCCTGCAACTGCTCGCGGAAACATCGATTTCTGGGGCGGGCTGCTTC	1529		
Db	1261	CTGCTGTGCGGGGCCCTGCAACTGCTCGCGGAAACATCGATTTCTGGGGCGGGCTGCTTC	1320		
Qy	1530	TTGGGACCTGAGGACTCTTTACAGACCCACAGAGACCCCAACCCACAGACTTGGGCCCCGGC	1589		
Db	1321	TTGGGACCTGAGGACTCTTTACAGACCCACAGAGACCCCAACCCACAGACTTGGGCCCCGGC	1380		
Qy	1590	CAGTGGCCATCAAAATCCGATGGAATTTGACGTGTGTCAGAGGAACCAATAAGATCAGG	1649		
Db	1381	CAGTGGCCATCAAAATCCGATGGAATTTGACGTGTGTCAGAGGAACCAATAAGATCAGG	1440		
Qy	1650	GACAGTGTGGGACCCCTGGCCGACGCGGACCCCTGACCGACCTTCGCTACGCTGTGG	1709		
Db	1441	GACAGTGTGGGACCCCTGGCCGACGCGGACCCCTGACCGACCTTCGCTACGCTGTGG	1500		
Qy	1710	GCGCGCTTCGTGTACCTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCTAGCGCGC	1769		
Db	1501	GCGCGCTTCGTGTACCTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCTAGCGCGC	1560		
Qy	1770	GCCAAACCCCGGGCGGCTCTACTGTGACAGATGCCCTATTCGCTGTATGTGTGACGAC	1829		
Db	1561	GCCAAACCCCGGGCGGCTCTACTGTGACAGATGCCCTATTCGCTGTATGTGTGACGAC	1620		
Qy	1830	GTGTTCTCTGCTGTGTGAGCGGCTGCGGCTCTTCTGACGCTGGCGCTGATCTTAC	1889		
Db	1621	GTGTTCTCTGCTGTGTGAGCGGCTGCGGCTCTTCTGACGCTGGCGCTGATCTTAC	1680		
Qy	1890	TCCGTGACACTGACAGTGAAGCGCTGTGTGGGAGAGAGAGAGCGCGCTGCGGACACC	1949		
Db	1681	TCCGTGACACTGACAGTGAAGCGCTGTGTGGGAGAGAGAGAGCGCGCTGCGGACACC	1740		
Qy	1950	ATGCGCGCAATGGGGCTCAGCGCGGCTGTCTGGGTAGGCTGCTTCTCAGTGTGCTC	2009		
Db	1741	ATGCGCGCAATGGGGCTCAGCGCGGCTGTCTGGGTAGGCTGCTTCTCAGTGTGCTC	1800		
Qy	2010	GGGCGCTTCTGCTCAGGCGCGCTGCTGTGCTTCTGCTCAAGCTGGGGGACATCTCTC	2069		
Db	1801	GGGCGCTTCTGCTCAGGCGCGCTGCTGTGCTTCTGCTCAAGCTGGGGGACATCTCTC	1860		
Qy	2070	CCCTACAGCAACCCCGGGCTGCTTCTGCTTCTTGGCAGCTTTCGCGGTGGCCACGCTG	2129		
Db	1861	CCCTACAGCAACCCCGGGCTGCTTCTGCTTCTTGGCAGCTTTCGCGGTGGCCACGCTG	1920		
Qy	2130	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTTGGCTGGCGCTGCGGC	2189		
Db	1921	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTTGGCTGGCGCTGCGGC	1980		
Qy	2190	GGCTTGGCTTCTTCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2249		
Db	1981	GGCTTGGCTTCTTCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040		
Qy	2250	CTGCCCGGGTGGCGCGCTGCGAGCTGCTGCTGCGCCGCTGGCTTGGCTTGGCTTGGCT	2309		
Db	2041	CTGCCCGGGTGGCGCGCTGCGAGCTGCTGCTGCGCCGCTGGCTTGGCTTGGCTTGGCT	2100		
Qy	2310	TGCGAGAGCTTGGCTTCTGCTGAGGAGCAGGGCGAGGGCGCGAGTGGCAACAGTGGGC	2369		

Db 2101 TCGAGAGCCTGGCTCTGTGAGGAGCAGGGCGCGAGTGGCACAAAGTGGGC 2160
Qy 2370 ACCGGGCTACGGCAGACGCTTTTCAGCCTGGCCAGGTCTCTGGCCTTCTGTGCTGGAC 2429
Db 2161 ACCGGGCTACGGCAGACGCTTTTCAGCCTGGCCAGGTCTCTGGCCTTCTGTGCTGGAC 2220
Qy 2430 GCGGCGCTCTACGGCCTTCGCCACTCTGTGTACTCTGAAGCTGTGTGCCAGGCCAGTAGGGG 2489
Db 2221 GCGGCGCTCTACGGCCTTCGCCACTCTGTGTACTCTGAAGCTGTGTGCCAGGCCAGTAGGGG 2280
Qy 2490 ATCCCTGAACCATGGAATTTTCTTTTCGAGGAGCTACTTGGTTCGGAGCTCTCGGCCCCCC 2549
Db 2281 ATCCCTGAACCATGGAATTTTCTTTTCGAGGAGCTACTTGGTTCGGAGCTCTCGGCCCCCC 2340
Qy 2550 AAGAGTCCAGCCCTTTCGCCCAACCCCGCTGGACCCCAAGGTGTGTAGAGGACACCG 2609
Db 2341 AAGAGTCCAGCCCTTTCGCCCAACCCCGCTGGACCCCAAGGTGTGTAGAGGACACCG 2400
Qy 2610 CCGGCGCTGAGTCTCTGGCGTATCCGTTTCGAGCTCTGGAGAGCGCTTTCTCTGAAGCCCG 2669
Db 2401 CCGGCGCTGAGTCTCTGGCGTCTCGGTTTCGAGCTCTGGAGAGCGCTTTCTCTGAAGCCCG 2460
Qy 2670 CAGCCAGCCTTCGGGGGCTCAGCCTGGACTTTCTACAGGGCCACATCACCGCCTTCCTG 2729
Db 2461 CAGCCAGCCTTCGGGGGCTCAGCCTGGACTTTCTACAGGGCCACATCACCGCCTTCCTG 2520
Qy 2730 GGCCACAAAGGGGCGGCAAGACCAACCCCTGTTCATCTTGTAGTGGCCTTTCCCAACC 2789
Db 2521 GGCCACAAAGGGGCGGCAAGACCAACCCCTGTTCATCTTGTAGTGGCCTTTCCCAACC 2580
Qy 2790 AGTGTGGCTCTGCTTCTATCTTGGGCGAGAGCTCGCTCCAGCATGGCGCGCATCCGG 2849
Db 2581 AGTGTGGCTCTGCTTCTATCTTGGGCGAGAGCTCGCTCCAGCATGGCGCGCATCCGG 2840
Qy 2850 CCCACCTGGGCGCTGTCTCTCAGTACAACGTGCTGTTTGACATGTGACCGTGGACGAG 2909
Db 2641 CCCACCTGGGCGCTGTCTCTCAGTACAACGTGCTGTTTGACATGTGACCGTGGACGAG 2700
Qy 2910 CAGTCTGGTTCTATGGGCGGCTGAAGGCTGTAGTGGCCTGTAGTGGGCGCGAGCAG 2969
Db 2701 CAGTCTGGTTCTATGGGCGGCTGAAGGCTGTAGTGGCCTGTAGTGGGCGCGAGCAG 2760
Qy 2970 GACCGTCTGTGAGGATGTGGGCTGGTCTCAGCAGAGTGTGACAGCTTCGCCACCTC 3029
Db 2761 GACCGTCTGTGAGGATGTGGGCTGGTCTCAGCAGAGTGTGACAGCTTCGCCACCTC 2820
Qy 3030 TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 3089
Db 2821 TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 2880
Qy 3090 GTTATCTGACGAGCCTACGGCTGGGCTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG 3149
Db 2881 GTTATCTGACGAGCCTACGGCTGGGCTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG 2940
Qy 3150 CTGCTGTCAAATAACGAGAGGTTCGACGCTGATCTCTCCACCCACACCTGGATGAG 3209
Db 2941 CTGCTGTCAAATAACGAGAGGTTCGACGCTGATCTCTCCACCCACACCTGGATGAG 3000
Qy 3210 GCAGAGTCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTGCTGTGGC 3269
Db 3001 GCAGAGTCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTGCTGTGGC 3060
Qy 3270 TCCCACTTCTTGGCGCGCTACCTGGGCTTCGGCTACTACTGAGCCTGGTGAAGGCC 3329
Db 3061 TCCCACTTCTTGGCGCGCTACCTGGGCTTCGGCTACTACTGAGCCTGGTGAAGGCC 3120
Qy 3330 CGCTGCTCTGACCAATGAGAGGCTGACACTGACATGAGAGGCGAGTGTGGACACC 3389
Db 3121 CGCTGCTCTGACCAATGAGAGGCTGACACTGACATGAGAGGCGAGTGTGGACACC 3180
Qy 3390 AGCAGGAAAAGAGATGGCAGCCAGGCGAGCAGTTCGGCACTCTCTCAGCTGCTGGCC 3449
Db 3181 AGCAGGAAAAGAGATGGCAGCCAGGCGAGCAGTTCGGCACTCTCTCAGCTGCTGGCC 3240

Qy 3450 CTGGTACGACACTGGGTGCCCGGGGCAACGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3509
Db 3241 CTGGTACGACACTGGGTGCCCGGGGCAACGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3300
Qy 3510 CTGGTCTGCTCCCTACACGCGGTGCCATGACGGCAGCTTCGCCACACTCTTTCGAGAGCTA 3569
Db 3301 CTGGTCTGCTCCCTACACGCGGTGCCATGACGGCAGCTTCGCCACACTCTTTCGAGAGCTA 3360
Qy 3570 GACACGCGGTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACGAGCTCGAG 3629
Db 3361 GACACGCGGTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACGAGCTCGAG 3420
Qy 3630 GAGATCTTCTGAAGGTGGTGGAGGAGTGTCTGCGGACACAGATATGAGGATGGCAGC 3689
Db 3421 GAGATCTTCTGAAGGTGGTGGAGGAGTGTCTGCGGACACAGATATGAGGATGGCAGC 3480
Qy 3690 TGGGGCAGCACCTATGACAGGCAATTGCTGGCCTAGACGTAAACCTCGGGCTCAAGATG 3749
Db 3481 TGGGGCAGCACCTATGACAGGCAATTGCTGGCCTAGACGTAAACCTCGGGCTCAAGATG 3540
Qy 3750 CGCCACAGGAGACAGCGCTGGAGAACGGGAAACGAGCTGGGTTCAGGCCACGAGACTGAC 3809
Db 3541 CGCCACAGGAGACAGCGCTGGAGAACGGGAAACGAGCTGGGTTCAGGCCACGAGACTGAC 3600
Qy 3810 CAGGGCTCTGGGCGACAGCGCTGGGCGGCTACAGGGCTGGGCACTGACCCGCCAGCAG 3869
Db 3601 CAGGGCTCTGGGCGACAGCGCTGGGCGGCTACAGGGCTGGGCACTGACCCGCCAGCAG 3660
Qy 3870 CTCAGGCGCTGTCTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGGCTGTTC 3929
Db 3661 CTCAGGCGCTGTCTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGGCTGTTC 3720
Qy 3930 GCCAGATGTGTGCTGCTGCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTG 3989
Db 3721 GCCAGATGTGTGCTGCTGCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTG 3780
Qy 3990 CTTCTTTTCGGGCACTACCGGCTCTGCGCTCAGTCCACCATGTACGCTGTCTCAGGTG 4049
Db 3781 CTTCTTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGTACGCTGTCTCAGGTG 3840
Qy 4050 TCTTCTTTCAGTGAGGACGCCCGAGGGGACCTTGGACGTGCCCGGTGTCTGAGGGCGTG 4109
Db 3841 TCTTCTTTCAGTGAGGACGCCCGAGGGGACCTTGGACGTGCCCGGTGTCTGAGGGCGTG 3900
Qy 4110 CTGACGAGCAGGACTGAGAGAGCCCCCAGTGCACGATAGTCCACACAGTTCCTGGCA 4169
Db 3901 CTGACGAGCAGGACTGAGAGAGCCCCCAGTGCACGATAGTCCACACAGTTCCTGGCA 3960
Qy 4170 CCAGAAAGTTCTCTGCTCAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4229
Db 3961 CCAGAAAGTTCTCTGCTCAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4020
Qy 4230 CCATCCCGAGCCTGCCAGTGTAGCAGCCCGGTGCCCGCGCTGTGCGCGACTGCGCG 4289
Db 4021 CCATCCCGAGCCTGCCAGTGTAGCAGCCCGGTGCCCGCGCTGTGCGCGACTGCGCG 4080
Qy 4290 GCTGAGCTGTGGTTCCTTCCCGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTCAG 4349
Db 4081 GCTGAGCTGTGGTTCCTTCCCGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTCAG 4140
Qy 4350 AACCTCAGAGGCCGGAACCTGTCTGACTTCTGCTCAAGACCTTACCCGCGCTTGGTGGC 4409
Db 4141 AACCTCAGAGGCCGGAACCTGTCTGACTTCTGCTCAAGACCTTACCCGCGCTTGGTGGC 4200
Qy 4410 CAGGGCTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACAGGAGCTTCTCGCTGGG 4469
Db 4201 CAGGGCTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACAGGAGCTTCTCGCTGGG 4260
Qy 4470 GGCCGAGACCCAGGCTTGCCTCTGGGCGCAAGAGTGGGCGGCTCAGTGGAGGATGGTGG 4529
Db 4261 GGCCGAGACCCAGGCTTGCCTCTGGGCGCAAGAGTGGGCGGCTCAGTGGAGGAGTGGTGG 4320

Qy	4530	CGCGTGTGAGTCCCGTCTGCGGGGGCCCTCGACCGTGTCTGTGAAAAACCTTCACAGCC	4589
Db	4321	CGCGTGTGAGTCCCGTCTGCGGGGGGGCCCTCGACCGTGTCTGTGAAAAACCTTCACAGCC	4380
Qy	4590	TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTTCAACAACAAGGCTGG	4649
Db	4381	TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTTCAACAACAAGGCTGG	4440
Qy	4650	CAC'TCCATGGTGGCTTTTGTCAACCGAGCCAGCAACGCAATCTCTCCGTGTCACTTGCCTC	4709
Db	4441	CAC'TCCATGGTGGCTTTTGTCAACCGAGCCAGCAACGCAATCTCTCCGTGTCACTTGCCTC	4500
Qy	4710	CCAGGCCGGCCGGCCAGGCCACAGCATACACACTCAACAACCCCTTGAACCTCAC	4769
Db	4501	CCAGGCCGGCCGGCCAGGCCACAGCATACACACTCAACAACCCCTTGAACCTCAC	4560
Qy	4770	AAGGAGCAGCTGTCTGAGGCTGCACTCATGGCCCTCTCGGTGAGCGTCTCGTCTCCATC	4829
Db	4561	AAGGAGCAGCTGTCTGAGGCTGCACTCATGGCCCTCTCGGTGAGCGTCTCGTCTCCATC	4620
Qy	4830	TGTTGGTCTTTTGGCCATGTCTTTTGTCTCCCGGCCAGCTTTCAC'TCTTGTTCCTCATTTAGGAG	4889
Db	4621	TGTTGGTCTTTTGGCCATGTCTTTTGTCTCCCGGCCAGCTTTCAC'TCTTGTTCCTCATTTAGGAG	4680
Qy	4890	CGAGTCAACCGAGCCAAAGCACTTGCACTCATGGGGGGCCCTGTCGCCCAACCCCTCTACTGG	4949
Db	4681	CGAGTCAACCGAGCCAAAGCACTTGCACTCATGGGGGGGGCCCTGTCGCCCAACCCCTCTACTGG	4740
Qy	4950	CTTGGCAACTTCTCTGGGACATGTGTAACTTACTTGGTGCACGATGCATCGTGGTGTCTC	5009
Db	4741	CTTGGCAACTTCTCTGGGACATGTGTAACTTACTTGGTGCACGATGCATCGTGGTGTCTC	4800
Qy	5010	ATCTTTCTGGCCCTTCCAGCAGAGGGCATA'TGTGGCCCTCGCCAACTGCGCTGTCTCTCCTG	5069
Db	4801	ATCTTTCTGGCCCTTCCAGCAGAGGGCATA'TGTGGCCCTCGCCAACTGCGCTGTCTCTCCTG	4860
Qy	5070	CTGTTGTCTACTATGTATGGCTGGTGCATCACACCGCTCAATGTAACCGAGCTCTCTTCTTC	5129
Db	4861	CTGTTGTCTACTATGTATGGCTGGTGCATCACACCGCTCAATGTAACCGAGCTCTCTTCTTC	4920
Qy	5130	TTCTCCGTGCCACAGCAGCTATGTGGTCTCACCTGCAATAACCTCTTTATTTGGCATC	5189
Db	4921	TTCTCCGTGCCACAGCAGCTATGTGGTCTCACCTGCAATAACCTCTTTATTTGGCATC	4980
Qy	5190	AATGGAAAGCATGGCCACTTTTGTCTTGAGCTCTTCTCTCATCAGAAGCTGCAGGAGGTG	5249
Db	4981	AATGGAAAGCATGGCCACTTTTGTCTTGAGCTCTTCTCTCATCAGAAGCTGCAGGAGGTG	5040
Qy	5250	AGCCGGATCTTGAAACAGGTCTTCTTATCTTCCCCACTTCTGCTTGGCCGGGGGCTC	5309
Db	5041	AGCCGGATCTTGAAACAGGTCTTCTTATCTTCCCCACTTCTGCTTGGCCGGGGGCTC	5100
Qy	5310	ATTGACATGTGGGAACGAGGCCATGGCTGATGCTTTTGAGCGCTTGGGAGACAGGCAG	5369
Db	5101	ATTGACATGTGGGAACGAGGCCATGGCTGATGCTTTTGAGCGCTTGGGAGACAGGCAG	5160
Qy	5370	TTCCAGTACCCCTGCGCTGGGAGGTGGTCGGCAAGAACTCTTTGGCCATGGTGATACAG	5429
Db	5161	TTCCAGTACCCCTGCGCTGGGAGGTGGTCGGCAAGAACTCTTTGGCCATGGTGATACAG	5220
Qy	5430	GGGCCCTCTTCTCTTCTTTCACACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAG	5489
Db	5221	GGGCCCTCTTCTCTTCTTTCACACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAG	5280
Qy	5490	CCAGGGTGAAGTCTCTGCCACTCTGGGAGAGGAGGACGAGATGTAGCCCGTGAACGG	5549
Db	5281	CCAGGGTGAAGTCTCTGCCACTCTGGGAGAGGAGGACGAGATGTAGCCCGTGAACGG	5340
Qy	5550	GAGCGGTGTCTCAAGGAGCCACCGAGGGGATGTGTGTGTCTGAGGAACTTGACCAAG	5609
Db	5341	GAGCGGTGTCTCAAGGAGCCACCGAGGGGATGTGTGTGTCTGAGGAACTTGACCAAG	5400
Qy	5610	GTATACGTGGGAGAGGATGCCAGCTGTTGACCGCTTGTGCTGGGGATTCCTCCCTGGT	5669

DB	5401	GTATACCGTGGCGACAGAGATGCCAGCTGTGTGACCGCTGTGCTCTGGGGATTCGCCCTCGGT	5460
QY	5670	GAGTGTTTTGGCGCTGCTGGGTGTGAATGAGCAGGGAAGACGTCACAGTTTTCGATGGTG	5729
DB	5461	GAGTGTTTTGGCGCTGCTGGGTGTGAATGAGCAGGGAAGACGTCACAGTTTTCGATGGTG	5520
QY	5730	ACGGGGACACATTTGGCCACAGCGGGCGAGGCTGTCTGCGAGGCCACACGCTGGCCCCGG	5789
DB	5521	ACGGGGACACATTTGGCCACAGCGGGCGAGGCTGTCTGCGAGGCCACACGCTGGCCCCGG	5580
QY	5790	GAACCCAGTCTGGGCACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAG	5849
DB	5581	GAACCCAGTCTGGGCACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAG	5640
QY	5850	CTGCTGACGGGCCGCGAGCACCTTGAGACTCTCTTTGCGCGCTGCGCGGTGTCCCGGAGGCC	5909
DB	5641	CTGCTGACGGGCCGCGAGCACCTTGAGACTCTCTTTGCGCGCTGCGCGGTGTCCCGGAGGCC	5700
QY	5910	CAGTTTCCCGACACCGCTGGCTCAGGCGCTGGCGCGTCTGSGACTCTCATGTGACGAGAC	5969
DB	5701	CAGTTTCCCGACACCGCTGGCTCAGGCGCTGGCGCGTCTGSGACTCTCATGTGACGAGAC	5760
QY	5970	CGGCGCTCGAGGACCTTACAGCGGAGGGAACAACGCAAGCTGCGGACGGCCCTTGGCGCTG	6029
DB	5761	CGGCGCTCGAGGACCTTACAGCGGAGGGAACAACGCAAGCTGCGGACGGCCCTTGGCGCTG	5820
QY	6030	GTTTGGGACCCAGCCGCTGGTCTTGAGCGAGCCGACCAACAGGCGATGAGCCGCCAGCGCG	6089
DB	5821	GTTTGGGACCCAGCCGCTGGTCTTGAGCGAGCCGACCAACAGGCGATGAGCCGCCAGCGCG	5980
QY	6090	CGGCGCTTCTTTGGAAACAGCCCTTTTGGCGCTGGTGGAGGCGCGCTTCAGTGATGCTC	6149
DB	5981	CGGCGCTTCTTTGGAAACAGCCCTTTTGGCGCTGGTGGAGGCGCGCTTCAGTGATGCTC	5940
QY	6150	ACCTCCCATAGCATGAGAGAGTGTGAAGCGCTCTGCTCGCGCTCAGCCATCATGTGTGAAT	6209
DB	5941	ACCTCCCATAGCATGAGAGAGTGTGAAGCGCTCTGCTCGCGCTCAGCCATCATGTGTGAAT	6000
QY	6210	GGGCGGTTCCGCTGCTGGGACGCCCGCAACATCTCAAGGGCAGATTCGCGGCGGCTCAC	6269
DB	6001	GGGCGGTTCCGCTGCTGGGACGCCCGCAACATCTCAAGGGCAGATTCGCGGCGGCTCAC	6060
QY	6270	ACACTGACCTTGGGGTGCCCGCCGCAAGTCTCCAGCCGCGCAGCGGCTTCGTGGCGGCG	6329
DB	6061	ACACTGACCTTGGGGTGCCCGCCGCAAGTCTCCAGCCGCGCAGCGGCTTCGTGGCGGCG	6120
QY	6330	GAGTTCCCTGGGTTCGAGAGCTGCGGAGGACATFGAGGCGCGCTGCGCTTCACAGCTGCCG	6389
DB	6121	GAGTTCCCTGGGTTCGAGAGCTGCGGAGGACATFGAGGCGCGCTGCGCTTCACAGCTGCCG	6180
QY	6390	CCGGGAGGGCGCTGCGCGCTGCGCGCTTTTGGAGAGCTGCGCGGTGACCGCGCGAGAG	6449
DB	6181	CCGGGAGGGCGCTGCGCGCTGCGCGCTTTTGGAGAGCTGCGCGGTGACCGCGCGAGAG	6240
QY	6450	CAGCGCTGGAGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTTC	6509
DB	6241	CAGCGCTGGAGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTTC	6300
QY	6510	TCCAAAGACCAAGGGGAAGCAGAGGACACCGAAGACGAAAGGAGGACGAGGTGGAGGTG	6569
DB	6301	TCCAAAGACCAAGGGGAAGCAGAGGACACCGAAGACGAAAGGAGGACGAGGTGGAGGTG	6360
QY	6570	GACCCCGCGGCGCTGACGACCCCAACCGGCTCAGCCAGTTCCTCCATGACCCCTAGC	6629
DB	6361	GACCCCGCGGCGCTGACGACCCCAACCGGCTCAGCCAGTTCCTCCATGACCCCTAGC	6420
QY	6630	ACTGCCGAGACTGTGCTCTGAGGCTCCCTCCCTCGCGGGGC---CGCGGGGAGGCCCTGG	6686
DB	6421	ACTGCCGAACTGTGCTCTGAGGCTCCCTCCCTTTCGCGGGCCCGCGGGGAGGCCCTGGG	6480
QY	6687	GAATGGCAAGGGCAAGGTAGAGTGTCTTAGAGCCCTGGACT	6727

Db	6481	GAATGGCAAGGGCAAGGTAAATGCTGTANGAGCCCTTGAAT	6521	
RESULT 8				
AX320364				
LOCUS	AX320364	6432 bp	DNA	linear PAT 14-DEC-2001
DEFINITION	Sequence 3 from Patent WO0187978.			
ACCESSION	AX320364			
VERSION	AX320364.1 GI:17901744			
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Glucksmann, M. and Curtis, R.A.			
TITLE	38594, a novel human transporter and uses thereof			
JOURNAL	Patent: WO 0187978-A 3 22-NOV-2001; Millennium Pharmaceuticals, Inc. (US)			
FEATURES	Location/Qualifiers			
source	1..6432			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
CDS	1..>6432			
	/notes="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAD19392.1"			
	/db_xref="GI:17901745"			
	/translation="WAFWTQLMLLKNFMVRRQPLLVELLWPLFFILVAVRHS			
	PPLHEHCFPNKPLPSAGTVPWLQGLI CNVNTCTFPQLTGPEEPGRSLNPNDSLVS			
	LIADARTVLHGASAHRTLAGLKGKLIATLRAARSTCAQPOPTKQSPLEPMLDVAELT			
	LSRTSLGALGAQOEPLHSLAEAGDLAQELLSLVELRALLOLRPGRTSGPFL			
	SEALCSVRGSPSTVGPLNWEASDLMELVQEPESALPDSSLPSPACSELIGALD			
	SHPLSRLLRUKPILQKLLFAPDTPFRKLMQVNRITFEELTLKRDVREWMELGPRIF			
	TFMNSSNVAMLRQLLQWDBGRQPRGPRGDHMEALRSLDPGSGSYWQDAHADV			
	HLVGTGVTECLSLDLKLEAAPSEAAVLSRALHLAHRFWAGVYFLGPDSDPT			
	HTPDLPGPHVRKIRIMDIDVVRTRNKIRDFWDPGPAADPLTLRYVMGCFVYLQ			
	DLV ERAAVLVSGANPRAGLVYLQMPYPCYVDVDFLVRSLRPLFTLTLAVLYSV			
	TLTKA VYREKETLRDTPMRAGLSRAVLMLGWFLSCLGPFLLSALLVLVLKGLDIL			
	PLYSHPG VVFLFLAFAVATVTSQFLLSAFFSRANLAACGLLAYFSLYFLVLCVAVRDL			
	LPAG GRVAFLSPVAFGSCSLALLEEQEGEAQWNVGTRPTADVFSLAQVGLLLD			
	AA LYGATLVLEAVLYPTVGTAGDSFATLRELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			
	KLKLSAAV GPEDRLQDVGLSVKQSVQTRHLSGGQKLSVAIAFVQSSQVILDE			
	TPAGDPA RGTWELLKLYREGRTLLSTLHDEALLGRVAVFVAGGLCCCGSP			
	FLRRHLSG YLTLVKARLPVLTNKKADTMESVDTRBKKNSQSGSRVGTPLLALVQHV			
	VPGAR LVLVEIPHELVLPVYTGADGSAFTRFELDTRLAELRLTGIGISDTLSLEI			
	FLKVV EACAQDMEGSGQHLCTAGLDVTLRKMPQETALENGEPAGSAPETDOGS			
	PD PGLSPGVSRLEKPFPGSPQALRGLSLDFYQGHITAFILHSGAGKTTLSIL			
	SLGLF PPSGSAFTLGHVDVSSMAAIRPHLGVCPQVNLVDFMLTVDEHVWFVGR			

Db 1075 |||||CAGCCAGACCTTGGAGGCGCGGACCAATGGAGGCGCTCGATCTCTTCTGGACCCCTGGG 1134
Qy 1350 AGCGGTGGCTACAGCTGGCAGACGACACACGCTGATGTGGGCACTCTGTGTGGGCAAGCTG 1409
Db 1135 AGCGGTGGCTACAGCTGGCAGACGACACACGCTGATGTGGGCACTCTGTGTGGGCAAGCTG 1194
Qy 1410 GCGGAGTGACGAGTGCTGCTGCTTGGGCAAGCTGGAGGCGGCACTCTCAGAGGACGCC 1469
Db 1195 GCGGAGTGACGAGTGCTGCTGCTTGGGCAAGCTGGAGGCGGCACTCTCAGAGGACGCC 1254
Qy 1470 CTGGGTGCGGGGCGCTGCAATGCTCGGGAACTGCAATCTGGGCGGCGCTGCTCTTC 1529
Db 1255 CTGGGTGCGGGGCGCTGCAATGCTCGGGAACTGCAATCTGGGCGGCGCTGCTCTTC 1314
Qy 1530 TTGGGACCTGAGGACTCTTTCAGACCCCAAGAGCACCCCAAGAGACCTGCGGCCCGGCG 1589
Db 1315 TTGGGACCTGAGGACTCTTTCAGACCCCAAGAGCACCCCAAGAGACCTGCGGCCCGGCG 1374
Qy 1590 CACGTGCGGATCAAAATCCGATGACATTTGACGTGGTCAAGAGGACCAATTAAGATCAGG 1649
Db 1375 CACGTGCGGATCAAAATCCGATGACATTTGACGTGGTCAAGAGGACCAATTAAGATCAGG 1434
Qy 1650 GACAGGTTTGGGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1709
Db 1435 GACAGGTTTGGGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1494
Qy 1710 GCGGCTTCTGTGTACTGCTGCAAGACCTGGTGAGCGTGACGCGGCTCGGCGTCTCAGCGGC 1769
Db 1495 GCGGCTTCTGTGTACTGCTGCAAGACCTGGTGAGCGTGACGCGGCTCGGCGTCTCAGCGGC 1554
Qy 1770 GCCAACCCCGGCGGCGGCTTACTGTCAGCAGATGCCCTATGCCGTATGTGACGAC 1829
Db 1555 GCCAACCCCGGCGGCGGCTTACTGTCAGCAGATGCCCTATGCCGTATGTGACGAC 1614
Qy 1830 GTGTTCTCTGTGTGCTGAGCGGCTGCTGCGGCTCTTCTGACGCTGGCTGGATCTAC 1889
Db 1615 GTGTTCTCTGTGTGCTGAGCGGCTGCTGCGGCTCTTCTGACGCTGGCTGGATCTAC 1674
Qy 1890 TCCGTGACACTGACAGTGAAGCCCTGTGCGGAGAGGAGACGCGGCTGCGGACACC 1949
Db 1675 TCCGTGACACTGACAGTGAAGCCCTGTGCGGAGAGGAGACGCGGCTGCGGACACC 1734
Qy 1950 ATGCGCGCATGGGCTCAGCGCGCGGTGCTGCGGCTAGGCTGGTCTCTCAGCTGCGCTC 2009
Db 1735 ATGCGCGCATGGGCTCAGCGCGCGGTGCTGCGGCTAGGCTGGTCTCTCAGCTGCGCTC 1794
Qy 2010 GGGCCCTTCTGCTCAGCGCGCGGTGCTGCTGCTGCTCAGCTGGGGAATCTCTC 2069
Db 1795 GGGCCCTTCTGCTCAGCGCGCGGTGCTGCTGCTGCTCAGCTGGGGAATCTCTC 1854
Qy 2070 CCTACAGCACCGGCGGTGCTGCTGCTTCTGGGAGGCTTCCGCGGTGGCCACCGGTG 2129
Db 1855 CCTACAGCACCGGCGGTGCTGCTGCTTCTGGGAGGCTTCCGCGGTGGCCACCGGTG 1914
Qy 2130 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGCGGCGCTGCGGC 2189
Db 1915 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGCGGCGCTGCGGC 1974
Qy 2190 GGCCTGGCTACTTCTCCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2249
Db 1975 GGCCTGGCTACTTCTCCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
Qy 2250 CTGCCCGGGGTGGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2309
Db 2035 CTGCCCGGGGTGGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
Qy 2310 TGGAGAGCTGGCTCTGCTGAGAGAGCAGGCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG 2369
Db 2095 TGGAGAGCTGGCTCTGCTGAGAGAGCAGGCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG 2154
Qy 2370 ACCCGGCTACGGCAGAGCTTCTACGCTGGCCAGAGTCTCTGCGCTTCTGCTGCTGCTGCTGCTG 2429

Db 2155 ACCCGGCTACGGCAGAGCTTCTTACGCTGGCCACAGGTCTCTGGCCTTCTGCTGCTGGAC 2214
Qy 2430 GCGGCGCTCTACGGCTCGCCACCTGTGTACCTGGAGCTGTGTGCCCGCAGGCACTGACGGG 2489
Db 2215 GCGGCGCTCTACGGCTCGCCACCTGTGTACCTGGAGCTGTGTGCCCGCAGGCACTGACGGG 2274
Qy 2490 ATCTCTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGTGTGGGAGCTCTCGGCGCCCG 2549
Db 2275 ATCTCTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGTGTGGGAGCTCTCGGCGCCCG 2334
Qy 2550 AAGAGTTCAGCCCTTGGCCCGCCCGCTGAGCCCAAGAGTCTGCTGGTAGAAGAGGACCG 2609
Db 2335 AAGAGTTCAGCCCTTGGCCCGCCCGCTGAGCCCAAGAGTCTGCTGGTAGAAGAGGACCG 2394
Qy 2610 CCGGCGCTGAGTCTGGGCTTCCGCTTCCGAGCTGGAGAGCGCTTTCTCTGAGAGCGCG 2669
Db 2395 CCGGCGCTGAGTCTGGGCTTCCGCTTCCGAGCTGGAGAGCGCTTTCTCTGAGAGCGCG 2454
Qy 2670 CAGCCAGCCCTGCGGGGCTCAGCCTGGAATTTCTACAGGGCCACATCACCGCCTTCTCTG 2729
Db 2455 CAGCCAGCCCTGCGGGGCTCAGCCTGGAATTTCTACAGGGCCACATCACCGCCTTCTCTG 2514
Qy 2730 GGCACAAACGGGCGCGGCAAGACCAACCTGCTCATCTTGTAGTGGCTCTTCCACCGC 2789
Db 2515 GGCACAGCGGGCGGCAAGACCAACCTGCTCATCTTGTAGTGGCTCTTCCACCGC 2574
Qy 2790 AGTGTGCTCTGCGCTTCCATCTGGGCGCAGCGCTCCGCTCCAGCATGGCGGCACTCGG 2849
Db 2575 AGTGTGCTCTGCGCTTCCATCTGGGCGCAGCGCTCCGCTCCAGCATGGCGGCACTCGG 2634
Qy 2850 CCCCACCTGGGCTGCTCTCAGTACAAACGTGCTTGTGTGACATGCTGACCGTGAGCGAG 2909
Db 2635 CCCCACCTGGGCTGCTCTCAGTACAAACGTGCTTGTGTGACATGCTGACCGTGAGCGAG 2694
Qy 2910 CAGCTGCTGCTATGGCGGCTGAAGGCTCTGAGTGGCGCTGTAGTGGGCGGCGGAGCGAG 2969
Db 2695 CAGCTGCTGCTATGGCGGCTGAAGGCTCTGAGTGGCGCTGTAGTGGGCGGCGGAGCGAG 2754
Qy 2970 GACGCTGCTGAGGATGTGGGCTGCTTCCAGAGCAGGTGTCAGACATGCGCCACCTC 3029
Db 2755 GACGCTGCTGAGGATGTGGGCTGCTTCCAGAGCAGGTGTCAGACATGCGCCACCTC 2814
Qy 3030 TCTGTGGGATGCAACCGGAAGCTGCTGCGTGCCCATGCTTGTGGGCGGCTTCCCAAGTT 3089
Db 2815 TCTGTGGGATGCAACCGGAAGCTGCTGCGTGCCCATGCTTGTGGGCGGCTTCCCAAGTT 2874
Qy 3090 GTTATCTGAGAGAGCTTACCGGCTGGCGTGAATCTGCTTCCGCGCGGCTTATTTGGAG 3149
Db 2875 GTTATCTGAGAGAGCTTACCGGCTGGCGTGAATCTGCTTCCGCGCGGCTTATTTGGAG 2934
Qy 3150 CTGCTGCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCCACACCTGGATGAG 3209
Db 2935 CTGCTGCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCCACACCTGGATGAG 2994
Qy 3210 GCAGAGCTGCTGGGAGACCGTGTGGCTGTGGCAGGTGGCGGCTTGTGTGCTGTGGC 3269
Db 2995 GCAGAGCTGCTGGGAGACCGTGTGGCTGTGGCAGGTGGCGGCTTGTGTGCTGTGGC 3054
Qy 3270 TCCCCACTTCTTCTGCGCGCTCAGCTGGGCTTCCGCGCTACTACTGACGCTGGTGAAGGCC 3329
Db 3055 TCCCCACTTCTTCTGCGCGCTCAGCTGGGCTTCCGCGCTACTACTGACGCTGGTGAAGGCC 3114
Qy 3330 GCGCTGCCCTGACCAACATGAGAGCTGACACTGACATGGAGGGCAGTGTGGACACC 3389
Db 3115 GCGCTGCCCTGACCAACATGAGAGCTGACACTGACATGGAGGGCAGTGTGGACACC 3174
Qy 3390 AGGCAGGAAGAAAGAAATGGCAGCGCAGGCACTGAGAGTGGCGACTCTCTCAGCTGTGGCC 3449
Db 3175 AGGCAGGAAGAAAGAAATGGCAGCGCAGGCACTGAGAGTGGCGACTCTCTCAGCTGTGGCC 3234
Qy 3450 CTGTGTACAGCACTGGGTGCGCGGCGCAGGCTGTGTGGAGAGCTGCGCACACGAGCTGTG 3509
Db 3235 CTGTGTACAGCACTGGGTGCGCGGCGCAGGCTGTGTGGAGAGCTGCGCACACGAGCTGTG 3294

Qy	3510	CTGGTGTGCGCCTACACGGGTGCCCATGACGGCAGCTTTCGCCACACTCTTCTCCGAGAGCTA	3569
Db	3295	CTGGTGTGCGCCTACACGGGTGCCCATGACGGCAGCTTTCGCCACACTCTTCTCCGAGAGCTA	3354
Qy	3570	GACACGGCGCTGGCGGAGCTGAGGCTCACTGGGCTACCGGATCTCCGACACCAAGCTTCGAG	3629
Db	3355	GACACGGCGCTGGCGGAGCTGAGGCTCACTGGGCTACCGGATCTCCGACACCAAGCTTCGAG	3414
Qy	3630	GAGATCTTCTGAAGTGTGGAGGAGTGTCTGCGGACACAGATATGGAGGATGGCAGC	3689
Db	3415	GAGATCTTCTGAAGTGTGGAGGAGTGTCTGCGGACACAGATATGGAGGATGGCAGC	3474
Qy	3690	TGCGGGCAGCACCTATGTCACAGGCATTGTGGGCTAGACGCTAAACCTTCGCGCTCAAGATG	3749
Db	3475	TGCGGGCAGCACCTATGTCACAGGCATTGTGGGCTAGACGCTAAACCTTCGCGCTCAAGATG	3534
Qy	3750	CCGCCACAGAGACAGCGCTGGAGAACCGGGAAACAGCTGGGTCAAGCCCCAGAGACTGAC	3809
Db	3535	CCGCCACAGAGACAGCGCTGGAGAACCGGGAAACAGCTGGGTCAAGCCCCAGAGACTGAC	3594
Qy	3810	CAGGGCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG	3869
Db	3595	CAGGGCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG	3654
Qy	3870	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGCCTGCCCGCGCAGCCGCGCGCTGTTC	3929
Db	3655	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGCCTGCCCGCGCAGCCGCGCGCTGTTC	3714
Qy	3930	GCCAGATCGTGCTGCTGCCCTCTTTTGTGGGCTTGCCCTCTGTTTCAGCCTCATCGTG	3989
Db	3715	GCCAGATCGTGCTGCTGCCCTCTTTTGTGGGCTTGCCCTCTGTTTCAGCCTCATCGTG	3774
Qy	3990	CTTCCTTTCCGGGCACCTACCGGCTCTCGGCTCAGTCCACCACTATACGCTGCTCAGGTC	4049
Db	3775	CTTCCTTTCCGGGCACCTACCGGCTCTCGGCTCAGTCCACCACTATACGCTGCTCAGGTC	3834
Qy	4050	TCCTTCTTCAGTGAGGACGCCCAAGGGACCTCTGAGCTGCCCGCTGCTCGAGGCGCTG	4109
Db	3835	TCCTTCTTCAGTGAGGACGCCCAAGGGACCTCTGAGCTGCCCGCTGCTCGAGGCGCTG	3894
Qy	4110	CTGCAGGACGAGACTGGAGGACCCCCAGTCAGTCAGCATAGCTCCACAGGTTCTCGGCA	4169
Db	3895	CTGCAGGACGAGACTGGAGGACCCCCAGTCAGTCAGCATAGCTCCACAGGTTCTCGGCA	3954
Qy	4170	CCAGAACTTCTGCTGAAGTGGCCCAAGGCTTTGGCCAGTGGCAACTGGACCCCGAGTCT	4229
Db	3955	CCAGAACTTCTGCTGAAGTGGCCCAAGGCTTTGGCCAGTGGCAACTGGACCCCGAGTCT	4014
Qy	4230	CCATCCCAGCTGCCAGTGTAGCCAGCCGGTSCCGGCGCTGCTGCGCGCTGCGGCTGCGG	4289
Db	4015	CCATCCCAGCTGCCAGTGTAGCCAGCCGGTSCCGGCGCTGCTGCGCGCTGCGGCTGCGG	4074
Qy	4290	GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGCACTGACCGGCTCTGGGGAAGTGGTTTTCAG	4349
Db	4075	GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGCACTGACCGGCTCTGGGGAAGTGGTTTTCAG	4134
Qy	4350	AACCTGACAGGCCGGGAACCTGTCTGACTTCTTGTGTAAGACCTTACCCGCGCTGGTGGCG	4409
Db	4135	AACCTGACAGGCCGGGAACCTGTCTGACTTCTTGTGTAAGACCTTACCCGCGCTGGTGGCG	4194
Qy	4410	CAGGGCTGTAAGACTTAAGAGTGGGTCAATGAGGTACAGGTACGGAGGCTTCTCGCTGGGG	4469
Db	4195	CAGGGCTGTAAGACTTAAGAGTGGGTCAATGAGGTACAGGTACGGAGGCTTCTCGCTGGGG	4254
Qy	4470	GGCCGAGACCCAGCGCTGCCCTCCGGGCCAAGAGTTGGGCCGCTCAGTGGAGGAGTTGTGG	4529
Db	4255	GGCCGAGACCCAGCGCTGCCCTCCGGGCCAAGAGTTGGGCCGCTCAGTGGAGGAGTTGTGG	4314
Qy	4530	GCGCTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTGTGAAAAAACCCTCACAGCC	4589
Db	4315	GCGCTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTGTGAAAAAACCCTCACAGCC	4374

Qy	4590	TGGGCTCACA	GGCTGANGTCT	CAGGACAGTCT	CAAGATCTGGT	TTCAACAACAAGGCTGG	4649
Db	4375	TGGGCTCACA	GGCTGANGTCT	CAGGACAGTCT	CAAGATCTGGT	TTCAACAACAAGGCTGG	4434
Qy	4650	CATCCCATGT	GGCTTTTGT	CAACCGAGCAG	CAACGCAATCT	CCGTGCTCACTGCCCC	4709
Db	4435	CATCCCATGT	GGCTTTTGT	CAACCGAGCAG	CAACGCAATCT	CCGTGCTCACTGCCCC	4494
Qy	4710	CCAGGCCCGC	CGCCACAGCC	CACAGCAT	CACACACT	CAACCAACCCCTTGAACCTCAC	4769
Db	4495	CCAGGCCCGC	CGCCACAGCC	CACAGCAT	CACACACT	CAACCAACCCCTTGAACCTCAC	4554
Qy	4770	AAGGAGCAG	CTGTCTGAGG	CTGCACTG	ATGAGGCTCT	CTCGGTGAGCGTCTCGTCTCCATC	4829
Db	4555	AAGGAGCAG	CTGTCTGAGG	CTGCACTG	ATGAGGCTCT	CTCGGTGAGCGTCTCGTCTCCATC	4614
Qy	4830	TGTGTGTCT	TTTGGCATG	TCTTTGTG	CCGGCCAGTCT	CACTCTTGCTCCATTTAGGAG	4889
Db	4615	TGTGTGTCT	TTTGGCATG	TCTTTGTG	CCGGCCAGTCT	CACTCTTGCTCCATTTAGGAG	4674
Qy	4890	CGAGTCAAC	CGAGCAGCA	CTTGACCT	CATGGGGGGC	TGTCGCCCAACCTCTACTCG	4949
Db	4675	CGAGTCAAC	CGAGCAGCA	CTTGACCT	CATGGGGGGC	TGTCGCCCAACCTCTACTCG	4734
Qy	4950	CTTGGCAACT	TTTCTTGGG	CATGTGTAA	CTACTTGGT	TGCCAGCATGCTGTGGTCTC	5009
Db	4735	CTTGGCAACT	TTTCTTGGG	CATGTGTAA	CTACTTGGT	TGCCAGCATGCTGTGGTCTC	4794
Qy	5010	ATCTTTCTG	GCCTTCAG	CAGAGGGC	ATATGTGG	CCCTTGCCTGCTCTCTCTG	5069
Db	4795	ATCTTTCTG	GCCTTCAG	CAGAGGGC	ATATGTGG	CCCTTGCCTGCTCTCTCTG	4854
Qy	5070	CTGTGTGCT	ACTACTGTAT	GTGCTGGT	CATCACACCG	CTCATGTATCCAGCGCTCCTTCTTC	5129
Db	4855	CTGTGTGCT	ACTACTGTAT	GTGCTGGT	CATCACACCG	CTCATGTATCCAGCGCTCCTTCTTC	4914
Qy	5130	TTCTCCGTG	CCCCAGC	ACAGCTAT	GTGGTCTC	CACTGCAATAAACCCTTTATTTGGCATC	5189
Db	4915	TTCTCCGTG	CCCCAGC	ACAGCTAT	GTGGTCTC	CACTGCAATAAACCCTTTATTTGGCATC	4974
Qy	5190	AATGGAAAG	CAATGGCC	ACTTTGTG	CTTGAGCT	CTTCTCTGATCAGAAGCTTGCAAGGAGTG	5249
Db	4975	AATGGAAAG	CAATGGCC	ACTTTGTG	CTTGAGCT	CTTCTCTGATCAGAAGCTTGCAAGGAGTG	5034
Qy	5250	AGCCGGATC	TTTGAACA	AGGCTCTT	CTTATCTT	CCCCACATCTTGCTTGGGCGCGGGGGCTC	5309
Db	5035	AGCCGGATC	TTTGAACA	AGGCTCTT	CTTATCTT	CCCCACATCTTGCTTGGGCGCGGGGGCTC	5094
Qy	5310	ATTGACATG	TGTCGGAA	CCAGGCCAT	CGCTGTAT	GCTCTTGAGCGCTTGGGAGACAGGCAG	5369
Db	5095	ATTGACATG	TGTCGGAA	CCAGGCCAT	CGCTGTAT	GCTCTTGAGCGCTTGGGAGACAGGCAG	5154
Qy	5370	TTCCAGTCA	CCCCCTG	CGCTGGG	AGTGCTG	GGCAAGAACCTCTTGCCCATGTGTATACAG	5429
Db	5155	TTCCAGTCA	CCCCCTG	CGCTGGG	AGTGCTG	GGCAAGAACCTCTTGCCCATGTGTATACAG	5214
Qy	5430	GGGCCCCCT	CTTCTCTT	CTTCACTAC	TACTGCTG	CAGCACCGAAGCCAACTCCTGCCACAG	5489
Db	5215	GGGCCCCCT	CTTCTCTT	CTTCACTAC	TACTGCTG	CAGCACCGAAGCCAACTCCTGCCACAG	5274
Qy	5490	CCAGGGTG	AGGTCTCT	CGCCACTC	CTCGGAGAGG	AGGACGAGGATGTAGCCCGGTGAACGG	5549
Db	5275	CCAGGGTG	AGGTCTCT	CGCCACTC	CTCGGAGAGG	AGGACGAGGATGTAGCCCGGTGAACGG	5334
Qy	5550	GAGCGGTGT	CCAAAGAG	CCACCCAG	GGGGATGTGT	TGTGCTGAGGAACTTTGACCAAG	5609
Db	5335	GAGCGGTGT	CCAAAGAG	CCACCCAG	GGGGATGTGT	TGTGCTGAGGAACTTTGACCAAG	5394
Qy	5610	GTATACCGT	GGGACAG	AGGATG	CCAGCTGTT	GACCGCTTGCTGGGGATTTCCCCCTGGT	5669
Db	5395	GTATACCGT	GGGACAG	AGGATG	CCAGCTGTT	GACCGCTTGCTGGGGATTTCCCCCTGGT	5454
Qy	5670	GAGTGT	TTTTGGG	CTGTGG	GTGTGA	ATGAGCAGGGAAGACGTCCAGTTTGGCATGGT	5729

Db	5455	GAGTGT TTGGCTCT GCTGGTGTGAAT GGAGCAGGGAAGACG TCCACGTT TCGCATGTG	5514
Qy	5730	ACGGGGGACACAT TGGCCAGCAGGCGCGAGG CTGTGCTGGCAGGCA CAGCGTGGCCCGG	5789
Db	5515	ACGGGGGACACAT TGGCCAGCAGGCGCGAGG CTGTGCTGGCAGGCA CAGCGTGGCCCGG	5574
Qy	5790	GAACCCAGTGTGCGC ACCTCAGCATGCGAT CTGCGCTCAAT CGATGCGCATCTTTGAG	5849
Db	5575	GAACCCAGTGTGCGC ACCTCAGCATGCGAT CTGCGCTCAAT CGATGCGCATCTTTGAG	5634
Qy	5850	CTGTGACGGGCGCG AGCAGCTGAGCTG CTTGGCGCGCTG CGCGGCTG CCCGAGGCC	5909
Db	5635	CTGTGACGGGCGCG AGCAGCTGAGCTG CTTGGCGCGCTG CGCGGCTG CCCGAGGCC	5694
Qy	5910	CAGGTGTCACAGC CGCTGGCTCAGCGCTG GGCGCTCTGGGACT CTCATGTACGCAGAC	5969
Db	5695	CAGGTGTCACAGC CGCTGGCTGCGGCTG GGGACTCTCATGTACGCAGAC	5754
Qy	5970	CGGCTCAGGACCT ACAGCGGAGGGA CAAAACGAA GCTGGCGAC CGGCTGCGCTG	6029
Db	5755	CGGCTCAGGACCT ACAGCGGAGGGA CAAAACGAA GCTGGCGAC CGGCTGCGCTG	5814
Qy	6030	GTGGGACCCAGC CGTGGTGTCTGACAGCGCC ACCAAGG CATGCA CCCGAGCGG	6089
Db	5815	GTGGGACCCAGC CGTGGTGTCTGACAGCGCC ACCAAGG CATGCA CCCGAGCGG	5874
Qy	6090	CGGCTTCTCTT TGAAACAGC CTTTGGCGGTG TGGCGGAGG CGCTTCAGTGTATGCTC	6149
Db	5875	CGGCTTCTCTT TGAAACAGC CTTTGGCGGTG TGGCGGAGG CGCTTCAGTGTATGCTC	5934
Qy	6150	ACCTCCCATAGC ATGAGGAGTGTGA GCGCTCTGCTCGCGCTAG CGCATCATG TGTAAT	6209
Db	5935	ACCTCCCATAGC ATGAGGAGTGTGA GCGCTCTGCTCGCGCTAG CGCATCATG TGTAAT	5994
Qy	6210	GGGCGGTTCCGCT GCTGGGAGCGCC CAACATCTCAAGG CAGATTGCGCGCGG GTAC	6269
Db	5995	GGGCGGTTCCGCT GCTGGGAGCGCC CAACATCTCAAGG CAGATTGCGCGCGG GTAC	6054
Qy	6270	ACACTGACCTT CGGCTGCGG TCCCAAGG TCCCAAGG TCCCAAGG TCCCAAGG TCCCAAGG	6329
Db	6055	ACACTGACCTT CGGCTGCGG TCCCAAGG TCCCAAGG TCCCAAGG TCCCAAGG TCCCAAGG	6114
Qy	6330	GAGTTCCTCGGT CGGAGCTGCGGAGG CACATGAGG CGGCTCGCTTCCAGCTGCGG	6389
Db	6115	GAGTTCCTCGGT CGGAGCTGCGGAGG CACATGAGG CGGCTCGCTTCCAGCTGCGG	6174
Qy	6390	CGGAGGCGGCT GCGGCGGCTCTTTGGAGAGCT GCGGCTGCA CGGCGCAGAG	6449
Db	6175	CGGAGGCGGCT GCGGCGGCTCTTTGGAGAGCT GCGGCTGCA CGGCGCAGAG	6234
Qy	6450	CACGCGTGGAGG ACTTTTCCGTGAGCCAGAC GATGCTGGAGGAGGTATTTCTTGACTTC	6509
Db	6235	CACGCGTGGAGG ACTTTTCCGTGAGCCAGAC GATGCTGGAGGAGGTATTTCTTGACTTC	6294
Qy	6510	TCCAAGACACAGG GAGGAGGAGC ACCGAAG CAGCAGGAGG GAGGAGGAGGAGG	6569
Db	6295	TCCAAGACACAGG GAGGAGGAGC ACCGAAG CAGCAGGAGG GAGGAGGAGGAGG	6354
Qy	6570	GACCCGCGCAGG CTGACGAC CCCAACCGCTGACCGAGTCTTCTCGATGAC CCCTAGC	6629
Db	6355	GACCCGCGCAGG CTGACGAC CCCAACCGCTGACCGAGTCTTCTCGATGAC CCCTAGC	6414
Qy	6630	ACTCCGAGACTGTGCTC	6647
Db	6415	ACTCCGAGACTGTGCTC	6432

RESULT 9
LOCUS AB055390
DEFINITION Homo sapiens ABCA7/ABCA-SSN mRNA for ABCA-SSN, complete cds.
ACCESSION AB055390

AB055390.1	GI:15042033	
Homo sapiens (human)		
Homo sapiens		
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
1		
Tanaka, A., Ikeda, Y., Abe-Dohmae, S., Arakawa, R., Sadanami, K., Kidera, A., Nakagawa, S., Nagase, T., Aoki, R., Kioka, N., Amachi, T., Yokoyama, S. and Ueda, K.		
Human ABCA1 contains a large amino-terminal extracellular domain homologous to an epitope of Sjogren's Syndrome		
Biochem. Biophys. Res. Commun. 283 (5), 1019-1025 (2001)		
21255283		
11355874		
2 (bases 1 to 6027)		
Ueda, K. and Kioka, N.		
Direct Submission		
Submitted (02-FEB-2001) Kazumitsu Ueda, Kyoto University, Division of Applied Life Sciences; Sakyu, Kyoto, Kyoto 606-8502, Japan (E-mail: ueda@kais.kyoto-u.ac.jp, Tel: 81-75-753-6105, Fax: 81-75-753-6104)		
Location/Qualifiers		
1..6027		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/chromosome="19"		
/cell_type="thymus"		
1..6027		
/gene="ABCA7/ABCA-SSN"		
1..6027		
/gene="ABCA7/ABCA-SSN"		
/codon_start=1		
/product="ABCA-SSN"		
/protein_id="BAB62294.1"		
/db_xref="GI:15042034"		
/translation="MVCLGTGQSGAGPLVSVQNHCPQGLSPQESLGLAGLQAGPELHS LLEAAEDLAQELALRLVELRALLQRPRTSGPLELLSEALCSVRGSPSTVGPLSNW YEADLMELVQBPESALPDSSLSLSPACSELIGALDSDHPLSLWRLKPLILGLKLLFA PDPTFKLMAQVNRTEELTLADRVREWMLGPRIETFNDSNVNMLORLQMD EGRQPQGRDRHMEALRSFLDPGSGTSDWADHADVGLVGTIGRVTESLQDLKLEA APESALVRSALQLLAEHRWAGVFLGPDSSDPTHEPTDLGPHVIRIKIMDIDV QMPVPCVDDVFLVLSRLPLFLTAWIYSVTLTKAVREKTRLDTRWAGLSL AVLWGLFSLCLGPFLLSALLVLKLDILPYSHQVWFLFLAAFAVATVQTSFLI SAFPSTRANLAAACGLAYFSLIPLYVLCVNRDRDPAGRVAASLLSPVAFGFCESL ALUEEQEGEGAHNVGTRPTADVFSLAQVSGLLLDLALYGLATWYLEAVCPQYGP EPWNFPFRSVCWCPRPKSPAPCTPDLPKVLEAPPGSPGVSRLSKRFPSPG QPALRGISLDFYQGHITAFILHNGAGKTTLSILSGLPFPGSGAFILGHVDRSMAA IRPHLGYCPOYNVLDMLTVDEHVMFYRLKGLSAAVVGPEODRLLODQVLSQSVQ TRHLSGQMORKLSVAIAFVGSQVILDEPTAGVDYDPSRSGIWEILLKYRGRTLIS THLDEALLGDRVAVVAGGRCLCCGSPFLRLRHLSGGYILTIVKARLPLTNEKADT DMGSDVTRQKNGSGSRVGTPTQLIALVQHWPVPGARLVLEELPHELVLVLPYTGAD IAGDVTFLRLMPQETALENGEPAGAPETDQSGDPDAGRVQGMALTRQOALLL KFLPLARRSRGLPAQIVLPAFLVGLALVFSLI VPRPHVPAIRLSPTMYCAQVSFFS EDAPGDPGRARLLLEALLQEALEPPVQHSRHPESAPEVPAEVAKVLASGNWTPES PAQCSQPGARRLLPDCPAAAGPPPPQAVTSGSEVVQNLITGRNLSDFLVKTYFLVR QGTRKWNVVRVYGGSLGGRDPLSPGELGSRVBEELWALLSPLPGGALDRVLKNL TAWHSLDQSLKIFNFWNWHSMVAFVNRASNAI LRAHLPGRARHSLITLNIHP LNTLKEQFPAALMASSVDVLCVVFVAMSFPASFTLVLISERVTRAKHLQMGGL SPITYLWGNFLWDMCNVLPACIWLVI FLAFOQRAYVAPANLPALLLLLLYKWSITP LMYPASFFSPSTAYVVLTCINLPIGINGSMATFVLELPSDKLQVSRLLKQVFLI FPHFLGRLIDWVRNQAMADAFERLGRQFQSPLRWEVVGKLLNLAWIQGLFELLFT LLLQHRSQLLPQPRVRLPLLGEEDVARERERVQGTQGVLDVLRNLTWVYKVRGQR APVADRLCLGI PFCECFGLGVNGAGKTSTFRMTGDTLASRGAVALSHGSAVREPSA AHLNMGYCPOSDAI FELTGREHLELLARLVPEAQVQVQTAGSGLARGLSWADRP AGTYSGNKRLKATALVGDPAVVFDEPTTGMDSARRFLWNLSLAVVREGSRVML TSHSMECEALCSRLAIWVNGRFLCSPQHLKGRFAAGHTLTLTVPAASQPAAPV AAFPGSELREAGGRURFQPPGRCALARVFGELAVHGAHEVDFSVFSQTMLEEV FLVFSKQDQKDETEQKEAGVGVDPAQLQHPKRVSQFLDDPSTAEVL"		

ORIGIN

Query Match 87.1%; Score 5928.2; DB 9; Length 6027;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 5936; Conservative 0; Mismatches 13; Indels 0; Gaps 0;									
Qy	702	CGCACGGAATCCCTGGGGTTGGCACTGGGGCCAAAGCCAGAGAGCCCTTTGCACAGCTTGTG	761						
Db	79	CCCCAGGAATCCCTGGGGTTGGCACTGGGGCCAAAGCCAGAGAGCCCTTTGCACAGCTTGTG	138						
Qy	762	GAGGCGCTGAGGACTTGGCCCCAGGAGCTCTGGCGCTGCGAGCCTGTGTGGAGCTTGG	821						
Db	139	GAGGCGCTGAGGACTTGGCCCCAGGAGCTCTGGCGCTGCGAGCCTGTGTGGAGCTTGG	198						
Qy	822	GCACTGTCAGAGACCCCGAGGAGCCAGCGGCCCTGTGAGTTGCTCAGAGGCCCTC	881						
Db	199	GCACTGTCAGAGACCCCGAGGAGCCAGCGGCCCTGTGAGTTGCTCAGAGGCCCTC	258						
Qy	882	TGCAGTGTACAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACGAGGCTAGT	941						
Db	259	TGCAGTGTACAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACGAGGCTAGT	318						
Qy	942	GACTGATGAGCTGTGGGGCAGGAGCCAGAAATTCGCCCTTGCACAGCAGCCTTGAGC	1001						
Db	319	GACTGATGAGCTGTGGGGCAGGAGCCAGAAATTCGCCCTTGCACAGCAGCCTTGAGC	378						
Qy	1002	CCGCGCTGCTCGAGCTGATTTGAGGACCTGGACAGCCAGCCGCTGTCCCGCTGCTCTGG	1061						
Db	379	CCGCGCTGCTCGAGCTGATTTGAGGACCTGGACAGCCAGCCGCTGTCCCGCTGCTCTGG	438						
Qy	1062	AGACGCTGAAGCTCTGTATCTCTGGGAAGCTACTCTTTTGCACAGATACACCTTTTACC	1121						
Db	439	AGACGCTGAAGCTCTGTATCTCTGGGAAGCTACTCTTTTGCACAGATACACCTTTTACC	498						
Qy	1122	CGGAAGCTCATGCCCCAGGTGAACCGGACCTTCGAGAGGCTCACCCCTGTGAGGGATGTC	1181						
Db	499	CGGAAGCTCATGCCCCAGGTGAACCGGACCTTCGAGAGGCTCACCCCTGTGAGGGATGTC	558						
Qy	1182	CGGAGGTGTGGAGATGCTTGGAGACCCCGATCTTACCTTCATGAACGACAGTTCCAAT	1241						
Db	559	CGGAGGTGTGGAGATGCTTGGAGACCCCGATCTTACCTTCATGAACGACAGTTCCAAT	618						
Qy	1242	GTGGCCATGTGACGCGGCTCCTGCAGATCAGATGAAGAAAGAGGAGCCACAGACCT	1301						
Db	619	GTGGCCATGTGACGCGGCTCCTGCAGATCAGATGAAGAAAGAGGAGCCACAGACCT	678						
Qy	1302	GGAGGCGGGAACCAATGGAGGCGCTTGCATCTTTCGACCTTGGAGCGGTGCGTAC	1361						
Db	679	GGAGGCGGGAACCAATGGAGGCGCTTGCATCTTTCGACCTTGGAGCGGTGCGTAC	738						
Qy	1362	AGCTGGCAGACGCAACGCTGATGTGGGCACTTGTGGGCACTGGGCGGAGTGACG	1421						
Db	739	AGCTGGCAGACGCAACGCTGATGTGGGCACTTGTGGGCACTGGGCGGAGTGACG	798						
Qy	1422	GAGTGCTGTCTTGGACAAAGCTGGAGCGGCACTTCAGAGGAGCGCCCTGTGTGCGGG	1481						
Db	799	GAGTGCTGTCTTGGACAAAGCTGGAGCGGCACTTCAGAGGAGCGCCCTGTGTGCGGG	858						
Qy	1482	GCCTTGCAATGTCTGCGGAACATCGATTTCTGGGCGGCGGTGCTTCTTTGGACCTTGAG	1541						
Db	859	GCCTTGCAATGTCTGCGGAACATCGATTTCTGGGCGGCGGTGCTTCTTTGGACCTTGAG	918						
Qy	1542	GACTCTTCAGACCCCAAGAGCACCCAAACCCAGACCTGGGCGCCCGGCACTGTGGCATC	1601						
Db	919	GACTCTTCAGACCCCAAGAGCACCCAAACCCAGACCTGGGCGCCCGGCACTGTGGCATC	978						
Qy	1602	AAATCCGATGACATTTGACGTGTTCACAGAGACCAATAAGATCAGGGACAGTTTGG	1661						
Db	979	AAATCCGATGACATTTGACGTGTTCACAGAGACCAATAAGATCAGGGACAGTTTGG	1038						
Qy	1662	GACCTGGCCAGCGCGGACCCCTTGACCGGACCTGCGCTACGTGTGGGCGGCTTCTGCTG	1721						
Db	1039	GACCTGGCCAGCGCGGACCCCTTGACCGGACCTGCGCTACGTGTGGGCGGCTTCTGCTG	1098						

Qy	1722	TACCTGCAAGACCTGCTGGAGCGTCCGCGTGTCTCAGCGCGCGCAACCCCGG	1781						
Db	1099	TACCTGCAAGACCTGCTGGAGCGTCCGCGTGTCTCAGCGCGCGCAACCCCGG	1158						
Qy	1782	GCGGCGCTTACCTGACGAGATGCCCTATPCGCTGTATGTGGACGAGTTCCTGCT	1841						
Db	1159	GCGGCGCTTACCTGACGAGATGCCCTATPCGCTGTATGTGGACGAGTTCCTGCT	1218						
Qy	1842	GTGCTAGCGCGTGTGCGGCTCTTCTTGAAGCTGCGCTTGAATCTACTCCGTGACACTG	1901						
Db	1219	GTGCTAGCGCGTGTGCGGCTCTTCTTGAAGCTGCGCTTGAATCTACTCCGTGACACTG	1278						
Qy	1902	ACAGTCAAGCGCGTGTGCGGAGAGAGAGACGCGCTGCGGAGACCACTACGCGCCATG	1961						
Db	1279	ACAGTCAAGCGCGTGTGCGGAGAGAGAGACGCGCTGCGGAGACCACTACGCGCCATG	1338						
Qy	1962	GCGCTCAGCGCGCGTGTCTGGCTAGGCTGTCTCAGCTGCTCGGGCGCTTCTCTG	2021						
Db	1339	GCGCTCAGCGCGCGTGTCTGGCTAGGCTGTCTCAGCTGCTCGGGCGCTTCTCTG	1398						
Qy	2022	CTCAGCGCGCGCTGTGCTTCTGGTCTCAAGCTGGGGAGACATCTCCCTACAGCCAC	2081						
Db	1399	CTCAGCGCGCGCTGTGCTTCTGGTCTCAAGCTGGGGAGACATCTCCCTACAGCCAC	1458						
Qy	2082	CGGGCGTGTCTTCTTCTTGGCAGCTTCCGCTGGCCACCGCTGACCCAGAGCTTC	2141						
Db	1459	CGGGCGTGTCTTCTTCTTGGCAGCTTCCGCTGGCCACCGCTGACCCAGAGCTTC	1518						
Qy	2142	CTGCTCAGCGCGCTTCTTCTCCGCGCAACTGGCTGCGGCTTGGCGCGCTTGGGCTAC	2201						
Db	1519	CTGCTCAGCGCGCTTCTTCTCCGCGCAACTGGCTGCGGCTTGGGCTTGGGCTAC	1578						
Qy	2202	TTCTCCCTTACCTGCGCTACGTGCTGTGTGGCTTGGGGGACCGGCTGCCCGGGT	2261						
Db	1579	TTCTCCCTTACCTGCGCTACGTGCTGTGTGGCTTGGGGGACCGGCTGCCCGGGT	1638						
Qy	2262	GGCGCGTGGCGGAGCTGCTGCTGCGCGCTTGGCTTCCGCTTGGCTGCGAGAGCTG	2321						
Db	1639	GGCGCGTGGCGGAGCTGCTGCTGCGCGCTTGGCTTCCGCTTGGCTGCGAGAGCTG	1698						
Qy	2322	GCTCTGCTGGAGGAGCAGGCGCGCAGTGGCAAACTGGGCAACCGTGGGCAACCGGCTACG	2381						
Db	1699	GCTCTGCTGGAGGAGCAGGCGCGCAGTGGCAAACTGGGCAACCGGCTTACG	1758						
Qy	2382	GCAGAGCTTCTTACGCTTGGCGGAGCTCTGCTGCTTCTGCTGCTGAGCGGCTTAC	2441						
Db	1759	GCAGAGCTTCTTACGCTTGGCGGAGCTCTGCTGCTTCTGCTGCTGAGCGGCTTAC	1818						
Qy	2442	GGCTCGCCACCTTGTACCTGGAAGCTGTGTGCGGCGGAGTACGGGATCTCCCTGAACCA	2501						
Db	1819	GGCTCGCCACCTTGTACCTGGAAGCTGTGTGCGGCGGAGTACGGGATCTCCCTGAACCA	1878						
Qy	2502	TGGAATTTTCTTTTCGAGGAGCTACTGCTGCGGACCTTGGCGCCCGCCCAAGAGTCCAGCC	2561						
Db	1879	TGGAATTTTCTTTTCGAGGAGCTACTGCTGCGGACCTTGGCGCCCGCCCAAGAGTCCAGCC	1938						
Qy	2562	CTTTGGCCCAACCCCGCTGAGCCAAAGGTGCTGTGAGAGAGGACCGCGCGGCTGAGT	2621						
Db	1939	CTTTGGCCCAACCCCGCTGAGCCAAAGGTGCTGTGAGAGAGGACCGCGCGGCTGAGT	1998						
Qy	2622	CTTTGGCGTATCCCTTCCGAGCTTGGAGAGCGCTTCTTCTGGAAGCCCGAGCGGCTG	2681						
Db	1999	CTTTGGCGTATCCCTTCCGAGCTTGGAGAGCGCTTCTTCTGGAAGCCCGAGCGGCTG	2058						
Qy	2682	CGGGGCTCAGCCTTGGACTTCTTACAGGCGGCAATCACCGCTTCTTGGGCGCAACCGG	2741						
Db	2059	CGGGGCTCAGCCTTGGACTTCTTACAGGCGGCAATCACCGCTTCTTGGGCGCAACCGG	2118						
Qy	2742	GCGGCAAGACCAACCTTGTCCATCTTGTGGCTCTTCCACCCAGTGTGGCTCT	2801						
Db	2119	GCGGCAAGACCAACCTTGTCCATCTTGTGGCTCTTCCACCCAGTGTGGCTCT	2178						
Qy	2802	GCCTTCTTGGGCGGACGAGCTCCGCTCCAGCATGGCGCCCATCGGCGCCACCTGGGC	2861						

Db 2179 |||||GCTTTCATCTCTGGGCGACAGGTCGCTCCAGCATGGCGGCATCCGCGCCCACTCGGGC 2238
Qy 2862 GTCTGCTCCTAGTACAAAGTGTGTTTGACATGCTGACCGGTGGACGACGACGTCTGGTTC 2921
Db 2239 GTCTGCTCCTAGTACAAAGTGTGTTTGACATGCTGACCGGTGGACGACGACGTCTGGTTC 2298
Qy 2922 TATGGGCGGCTGAAGGGTCTGAGTGCCTGTGAGTGGGCCCCGAGACGACGACGTCGTCTGCTG 2981
Db 2299 TATGGGCGGCTGAAGGGTCTGAGTGCCTGTGAGTGGGCCCCGAGACGACGACGTCGTCTGCTG 2358
Qy 2982 CAGGATGTGGGGCTGCTTCCAAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTTGGGATG 3041
Db 2359 CAGGATGTGGGGCTGCTTCCAAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTTGGGATG 2418
Qy 3042 CAAGGAAGCTGTCCGTGGCCATTGCCCTTGTGGCGGCTCCCAAGTTGTTATCTCTGGAC 3101
Db 2419 CAAGGAAGCTGTCCGTGGCCATTGCCCTTGTGGCGGCTCCCAAGTTGTTATCTCTGGAC 2478
Qy 3102 GAGCTACGGCTGGCGTGGATCTCTGCTTCCCGCGCGGTATTTGGGAGCTGCTGCTCAAA 3161
Db 2479 GAGCTACGGCTGGCGTGGATCTCTGCTTCCCGCGCGGTATTTGGGAGCTGCTGCTCAAA 2538
Qy 3162 TACCGAGAGCTCGCACGCTGATCTCTCCACCCACACCTGGATGGACGACGAGCTGCTG 3221
Db 2539 TACCGAGAGCTCGCACGCTGATCTCTCCACCCACACCTGGATGGACGACGAGCTGCTG 2598
Qy 3222 GGAGACCTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 3281
Db 2599 GGAGACCTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 2658
Qy 3282 CTGCGCGCTACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCGCGCTGCCCCTG 3341
Db 2659 CTGCGCGCTACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCGCGCTGCCCCTG 2718
Qy 3342 ACCACCAATGAGAGGCTGACACTGACATGAGGGCAGTGTGGACACAGGACGAGAAAG 3401
Db 2719 ACCACCAATGAGAGGCTGACACTGACATGAGGGCAGTGTGGACACAGGACGAGAAAG 2778
Qy 3402 AAGAAATGACGACGAGGACAGAGTGGGCACTCTCAGCTGCTGGCCCTGATACAGCAC 3461
Db 2779 AAGAAATGACGACGAGGACAGAGTGGGCACTCTCAGCTGCTGGCCCTGATACAGCAC 2838
Qy 3462 TGGGTGCCCGGGGACCGGCTGGTGGAGAGCTGCCACACGAGCTGGTGTGCTGCGCC 3521
Db 2839 TGGGTGCCCGGGGACCGGCTGGTGGAGAGCTGCCACACGAGCTGGTGTGCTGCGCC 2898
Qy 3522 TACAGGGGTGCCATGACGGCAGCTTGCACACACTCTTCCGAGAGCTAGACACGGGGCTG 3581
Db 2899 TACAGGGGTGCCATGACGGCAGCTTGCACACACTCTTCCGAGAGCTAGACACGGGGCTG 2958
Qy 3582 GCGGAGCTGAGGCTCACTGGCTAGCGGATCTCCGACACACGCTCGAGGAGATCTTCCCTG 3641
Db 2959 GCGGAGCTGAGGCTCACTGGCTAGCGGATCTCCGACACACGCTCGAGGAGATCTTCCCTG 3018
Qy 3642 AAGTGTGTGAGGAGTGTGCTGCGGACACAGATATGAGGATGGCAGCTGCGGGCAGCAC 3701
Db 3019 AAGTGTGTGAGGAGTGTGCTGCGGACACAGATATGAGGATGGCAGCTGCGGGCAGCAC 3078
Qy 3702 CTATGACACAGGCAATGCTGGCTAGACGTAAACCTTGGGCTCAAGATGCCGCCACAGAG 3761
Db 3079 CTATGACACAGGCAATGCTGGCTAGACGTAAACCTTGGGCTCAAGATGCCGCCACAGAG 3138
Qy 3762 ACAGGCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCAGAGACTGACCGGGCTCTGGG 3821
Db 3139 ACAGGCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCAGAGACTGACCGGGCTCTGGG 3198
Qy 3822 CCAGACGCGTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGGCCCTG 3881
Db 3199 CCAGACGCGTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGGCCCTG 3258
Qy 3882 CTTCTCAAGCGCTTCTGCTTCCCGCGCGCAGCGCGCGCTGTTCCGCCAGATCGTG 3941

Db 3259 CTTCTCAAGCGCTTCTGCTTCCCGCGCGCAGCGCGCGCTGTTCCGCCAGATCGTG 3318
Qy 3942 CTGCTGCCCTCTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTGCCTCTTTTCGGG 4001
Db 3319 CTGCTGCCCTCTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTGCCTCTTTTCGGG 3378
Qy 4002 CACTACCGGCTCTGCGGCTCAGTCCCAACCATGATGACGATGCTCAGGTGTCTTTCAGT 4061
Db 3379 CACTACCGGCTCTGCGGCTCAGTCCCAACCATGATGACGATGCTCAGGTGTCTTTCAGT 3438
Qy 4062 GAGACGCGCCAGAGGGACCTCGAGCGTCCCGGCTGCTCGAGGGCTGCTCGAGGAGCA 4121
Db 3439 GAGACGCGCCAGAGGGACCTCGAGCGTCCCGGCTGCTCGAGGGCTGCTCGAGGAGCA 3498
Qy 4122 GGACTGAGAGAGCCCCCAGTGCAGATAGTCCCAAGTTCCTGGGACACCAAGTTCCT 4181
Db 3499 GGACTGAGAGAGCCCCCAGTGCAGATAGTCCCAAGTTCCTGGGACACCAAGTTCCT 3558
Qy 4182 GCTGAAGTGGCAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCCAGCC 4241
Db 3559 GCTGAAGTGGCAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCCAGCC 3618
Qy 4242 TGCCAGTGTAGCCAGCGGCTGCCCGGCTGCTGCCGACTGCTGCCGCTGCACTGGT 4301
Db 3619 TGCCAGTGTAGCCAGCGGCTGCCCGGCTGCTGCCGACTGCTGCCGCTGCACTGGT 3678
Qy 4302 GGTCCCCCTCGGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTCAGAACCTGCAGAGC 4361
Db 3679 GGTCCCCCTCGGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTCAGAACCTGCAGAGC 3738
Qy 4362 CGGAACCTGTCTGACTTCTGGTCAAGACCTTACCGCGCTGGTGGCGCCAGGGCTGAAG 4421
Db 3739 CGGAACCTGTCTGACTTCTGGTCAAGACCTTACCGCGCTGGTGGCGCCAGGGCTGAAG 3798
Qy 4422 ACTAAGAAGTGGGTGAATGAGGTCAAGGTACCGAGGCTTCTCGCTGGGGGGCGGAGACCA 4481
Db 3799 ACTAAGAAGTGGGTGAATGAGGTCAAGGTACCGAGGCTTCTCGCTGGGGGGCGGAGACCA 3858
Qy 4482 GGCCTGCCCTCGGCCCAAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGT 4541
Db 3859 GGCCTGCCCTCGGCCCAAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGT 3918
Qy 4542 CCCCTGCTGCGGGGCGCTCGACCGTGTCTGAAAAAACCCTCACAGCTGGGCTCACAGC 4601
Db 3919 CCCCTGCTGCGGGGCGCTCGACCGTGTCTGAAAAAACCCTCACAGCTGGGCTCACAGC 3978
Qy 4602 CTGATGCTCAGGACAGTCTCAAGATCTGGTTCACAAAGGCTGGCACTCCATGCTG 4661
Db 3979 CTGATGCTCAGGACAGTCTCAAGATCTGGTTCACAAAGGCTGGCACTCCATGCTG 4038
Qy 4662 GCTTTGTCAACCGAGCGCAGCAATCCTCGTGTCTCACTGCGCTCACTGCGCCCGAGGCGCGGC 4721
Db 4039 GCTTTGTCAACCGAGCGCAGCAATCCTCGTGTCTCACTGCGCTCACTGCGCCCGAGGCGCGGC 4098
Qy 4722 CGCACGCCCAAGCATCACCACTCAACCAACCCCTTGAACCTCACCAAGGAGCAGCTG 4781
Db 4099 CGCACGCCCAAGCATCACCACTCAACCAACCCCTTGAACCTCACCAAGGAGCAGCTG 4158
Qy 4782 TCTGAGGCTGCACTGATGGGCTCTCGGTGAGCGTCTCGTCTCCATCTGCTGTGCTCTT 4841
Db 4159 TTTGAGGCTGCACTGATGGGCTCTCGGTGAGCGTCTCGTCTCCATCTGCTGTGCTCTT 4218
Qy 4842 GCAATGCTCTTGTCCCGGCGCAGCTTCACTCTGTCTCTCACTTGAAGGAGCAGTCAACCGA 4901
Db 4219 GCAATGCTCTTGTCCCGGCGCAGCTTCACTCTGTCTCTCACTTGAAGGAGCAGTCAACCGA 4278
Qy 4902 GCCAAGCAGCTGAGCTCATGGGGGCGCTGTCCCCCAACCCCTCTACTGCTGCTTGGCACTTT 4961
Db 4279 GCCAAGCAGCTGAGCTCATGGGGGCGCTGTCCCCCAACCCCTCTACTGCTGCTTGGCACTTT 4338
Qy 4962 CTCTGGGACATGTGTAACTACTTGGTGCAGCATGCACTGCTGTGCTCATCTTTCTGGCC 5021
Db 4339 CTCTGGGACATGTGTAACTACTTGGTGCAGCATGCACTGCTGTGCTCATCTTTCTGGCC 4398

Qy	5022	TTCCAGCAGAGGGCATATGTGGCCCTGCCAACCTGCCTCTCTCTGCTGTTGCTACTA	5081
Db	4399	TTCCAGCAGAGGGCATATGTGGCCCTGCCAACCTGCCTCTCTCTGCTGTTGCTACTA	4458
Qy	5082	CTGTATGGCTGGTGCATCAACACGCTCATGTATCCAGCCTCTCTTCTTCTCCGTGCCCC	5141
Db	4459	CTGTATGGCTGGTGCATCAACACGCTCATGTATCCAGCCTCTCTTCTTCTCCGTGCCCC	4518
Qy	5142	AGCACGCCATATGGTGGCTCACCTGCATAAACCTCTTTATTTGGCATCAATGAAGCATG	5201
Db	4519	AGCACGCCATATGGTGGCTCACCTGCATAAACCTCTTTATTTGGCATCAATGAAGCATG	4578
Qy	5202	GCCACCTTTGTGCTTGAGCTCTTCTCTGTATCAGAAGCTGCAGGAGTGAACCGGATCTTG	5261
Db	4579	GCCACCTTTGTGCTTGAGCTCTTCTCTGTATCAGAAGCTGCAGGAGTGAACCGGATCTTG	4638
Qy	5262	AAACAGGCTTTCCTTATCTTCCCCCACTTCTGCTTGGGCCGGGGCTCATTTGACATGGTG	5321
Db	4639	AAACAGGCTTTCCTTATCTTCCCCCACTTCTGCTTGGGCCGGGGCTTATTTGACATGGTG	4698
Qy	5322	CGNAACCAGGCCATGGCTGTATGCTTTGAGGCTTTGGGACACAGGCAGTTCCAGTCACCC	5381
Db	4699	CGNAACCAGGCCATGGCTGTATGCTTTGAGGCTTTGGGACACAGGCAGTTTCCAGTCACCC	4758
Qy	5382	CTGGCTTGGAGGTGGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTCTTC	5441
Db	4759	CTGGCTTGGAGGTGGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTCTTC	4818
Qy	5442	CTTCTCTTCACTACTGCTGCGACACCGAAGCCAACTCTCTGCCACACGCCAGCGGTGAGG	5501
Db	4819	CTTCTCTTCACTACTGCTGCGACACCGAAGCCAACTCTCTGCCACACGCCAGCGGTGAGG	4878
Qy	5502	TCTCTGCCACTCTGGAGAGGAGGACGAGATGTAGCCCGTGAAACGGAGACGGGTGGTC	5561
Db	4879	TCTCTGCCACTCTGGAGAGGAGGACGAGATGTAGCCCGTGAAACGGAGACGGGTGGTC	4938
Qy	5562	CAAGGACCCACCAAGGGGATGTGTGGTCTGAGGAACCTTGACCAAGGTATACCGTGGG	5621
Db	4939	CAAGGACCCACCAAGGGGATGTGTGGTCTGAGGAACCTTGACCAAGGTATACCGTGGG	4998
Qy	5622	CAGAGGATGCCAGCTGTGTGACCGCTTGTGCTGGGGATTCGCCCTGCTGAGTGTTTGGG	5681
Db	4999	CAGAGGATGCCAGCTGTGTGACCGCTTGTGCTGGGGATTCGCCCTGCTGAGTGTTTGGG	5058
Qy	5682	CTGCTGGGTGTGAATGAGACAGGGAACAGCTCAACGTTTCGATGTGTGACGGGGGACACA	5741
Db	5059	CTGCTGGGTGTGAATGAGACAGGGAACAGCTCAACGTTTCGATGTGTGACGGGGGACACA	5118
Qy	5742	TTGGCCACAGGGGCGAGGCTGTCTCGGACGGCCACAGCGTGGCCGGGAACCCAGTGCT	5801
Db	5119	TTGGCCACAGGGGCGAGGCTGTCTCGGACGGCCACAGCGTGGCCGGGAACCCAGTGCT	5178
Qy	5802	GGCACCCTCAGCATGGGATATCTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5861
Db	5179	GGCACCCTCAGCATGGGATATCTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5238
Qy	5862	CGCAGACACTGTGAGCTGTCTTGGCGCTTGGCGGTTGTCCCGAGGGCCCAAGTTGCCAG	5921
Db	5239	CGCAGACACTGTGAGCTGTCTTGGCGCTTGGCGGTTGTCCCGAGGGCCCAAGTTGCCAG	5298
Qy	5922	ACCGCTGGCTCAGCCCTGGCGCTCTGAGGACTCTCATGTGTACGACACCGCCTGCAGGC	5981
Db	5299	ACCGCTGGCTCAGCCCTGGCGCTCTGAGGACTCTCATGTGTACGACACCGCCTGCAGGC	5358
Qy	5982	ACCTACAGCGAGGAGAAACAAACGCAAGCTGGCGACCGCCCTGGCGCTGGTTGGGGACCCA	6041
Db	5359	ACCTACAGCGAGGAGAAACAAACGCAAGCTGGCGACCGCCCTGGCGCTGGTTGGGGACCCA	5418
Qy	6042	GCCGTGGTGTCTTGAGACGACCGAACACAGGCATGAAACCCACGCGCGCGCGCTTCCTT	6101
Db	5419	GCCGTGGTGTCTTGAGACGACCGAACACAGGCATGAAACCCACGCGCGCGCGCTTCCTT	5478

Qy	6102	TGAAACAGACCTTTTGGCCGTGTGTGCGGAGGGCCGTTTCACTGATGTCTCACTCCCATAGC	6161
Db	5479	TGAAACAGACCTTTTGGCCGTGTGTGCGGAGGGCCGTTTCACTGATGTCTCACTCCCATAGC	5538
Qy	6162	ATGGAGGAGTGTGAAGCGCTCTGCTCGCCCTAGCCATCATGTGTGAATGGCGGTTCCTCCG	6221
Db	5539	ATGGAGGAGTGTGAAGCGCTCTGCTCGCCCTAGCCATCATGTGTGAATGGCGGTTCCTCCG	5598
Qy	6222	TGCTTGGCAGCCCGCAACATCTCAAGGGCAGATTTCGCGCGGGTACACACTGACCCCTG	6281
Db	5599	TGCTTGGCAGCCCGCAACATCTCAAGGGCAGATTTCGCGCGGGTACACACTGACCCCTG	5658
Qy	6282	CGGTGTCCTCCCGCAAGGTCCTCCAGCCGCGCAGCGCCCTTCGTGCGCGCGAGTTTCCCTGGG	6341
Db	5659	CGGTGTCCTCCCGCAAGGTCCTCCAGCCGCGCAGCGCCCTTCGTGCGCGCGAGTTTCCCTGGG	5718
Qy	6342	TGCGAGCTTGGCGAGGCAATGAGAGCGCCTCGCGTTTCAGCTGCGCGCGGAGAGGCGC	6401
Db	5719	TGCGAGCTTGGCGAGGCAATGAGAGCGCCTCGCGTTTCAGCTGCGCGCGGAGAGGCGC	5778
Qy	6402	TGCGCCTTCGCGCGCGTCTTTTGGAGAGCTGGCGGTGCACGGCGCAGAGCACGGCGTGGAG	6461
Db	5779	TGCGCCTTCGCGCGCGTCTTTTGGAGAGCTGGCGGTGCACGGCGCAGAGCACGGCGTGGAG	5838
Qy	6462	GACTTTTTCGTGAGCCAGACGATGCTGGAGAGGTATTCTTTGACTTCTCCAAAGGACCA	6521
Db	5839	GACTTTTTCGTGAGCCAGACGATGCTGGAGAGGTATTCTTTGACTTCTCCAAAGGACCA	5898
Qy	6522	GGAGAGCAGAGACACCGAAGACGACGAGGAGGAGGAGTGGAGTGGACCCCGCGCCA	6581
Db	5899	GGAGAGCAGAGACACCGAAGACGACGAGGAGGAGGAGTGGAGTGGACCCCGCGCCA	5958
Qy	6582	GGCTCGACGACCCCAAAACGCTCAGCCAGTCTCTCGATCACCTAGCACTGCCGAGACT	6641
Db	5959	GGCTCGACGACCCCAAAACGCTCAGCCAGTCTCTCGATCACCTAGCACTGCCGAGACT	6018
Qy	6642	GTGCTCTGA 6650	
Db	6019	GTGCTCTGA 6027	
RESULT 10			
LOCUS	CQ719367	6324 bp	DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 5301 from Patent WO02068579.		
ACCESSION	CQ719367		
VERSION	CQ719367.1	GI:42280224	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 5301 06-SEP-2002;		
PE Corporation (NY) (US)			
FEATURES	Location/Qualifiers		
source	1..6324		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	81.7%;	Score 5557.2;	DB 6; Length 6324;
Best Local Similarity	92.9%;	Pred. No. 0;	
Matches 6047; Conservative	0; Mismatches 208;	Indels 255; Gaps 11;	
Qy	210	ATGCCCTTCGGACACAGCTGATGCTGCTCTGGAAGAAATTTCACTATCGCCGAGA	269
Db	1	ATGCCCTTCGGACACAGCTGATGCTGCTCTGGAAGAAATTTCACTATCGCCGAGA	60

Db 2092 GCGGCGCTCTACGGCTCGCCACCTGGTACTTGGAGCTGTGTGCCAGGCACTAGTACGG 2151
QY 2490 ATCCCTGAACCATGGAATTTTCTTTTCGAGAGCTACTGTGTGGGACCTCGGCCCCC 2549
Db 2152 ATCCCTGAACCATGGAATTTTCTTTTCGAGAGCTACTGTGTGGGACCTCGGCCCCC 2211
QY 2550 AAGAGTCCAGGCTTGGCCCAACCCGCTGGACCCAAAGGTGCTGTAGAGGACCG 2609
Db 2212 AAGAGTCCAGGCTTGGCCCAACCCGCTGGACCCAAAGGTGAGGCCCTACAGAGCTTA 2271
QY 2610 CCGGCGCTGAGTCTGCGGTATCCGTTTCGAGCTGTGGAGAGCGCTTTCTCGAAGCCCC 2669
Db 2272 ATAGCTGTTGTCCACATATGCCAGT-TCCCCCCCAACAGAGATTCACAGATGCCAA 2330
QY 2670 CAGCCAGCTTCGGGGCTCAGCTGGAATTTCTACAGGGCCACATACCGCTTCCTG 2729
Db 2331 TGACAAATGACTGGACACCCCAACCTCTCACACCTGCCCTGAAGACACTGCGGTCCCCAAG 2390
QY 2730 GGCCACAAAGGGCCGCGCAGACCAACACCTGTCCATCTTGAGTGGCCTCTTCCACCC 2789
Db 2391 CTC---CGCAGCTTTATAGGCCCGCGCCAGCAGCTGGAGAGCGCTTTCTTGGA 2446
QY 2790 AGTGTGGCTCTGCTTATCTCTGGGCCACGAGCTCGCTCCAGCATGGCGCCATCCGG 2849
Db 2447 GCCCGCAGC---CAGCCCTGGGGGCTCAGCTGGACTTCTACAGGGCCACATCACC 2502
QY 2850 CCGCACTGTGGCTGTCTCTCAGTACAAACGTGTCTTTGACATGTGACCGTGGACGAG 2909
Db 2503 GCCTTCTGGGCCAC---AACGGGGCCGCGCAAGACCAACACCTGTGACCGTGGACGAG 2559
QY 2910 CAGTCTGTCTTATGGCGGCTGAGGGTCTGAGTCCGCTGTAGTGGGCCCGGAGCAG 2969
Db 2560 CACGTCTGTCTTATGGCGGCTGAGGGTCTGAGTCCGCTGTAGTGGGCCCGGAGCAG 2619
QY 2970 GACCGTCTGTGACAGATGTGGGCTGTCTCAAGCAGAGTGTGACAGTCTCGCCACCTC 3029
Db 2620 GACCGTCTGTGACAGATGTGGGCTGTCTCAAGCAGAGTGTGACAGTCTCGCCACCTC 2679
QY 3030 TCTGTGGGATGCAACGGAAAGCTGTCTGTGGCCATTTGCTGTGGCGGCTCCCAAGTT 3089
Db 2680 TCTGTGGGATGCAACGGAAAGCTGTCTGTGGCCATTTGCTGTGGCGGCTCCCAAGTT 2739
QY 3090 GTTATCTGACGAGCTACGCTGGCTGGATCTGCTTCCCGCGCGGTATTTGGAG 3149
Db 2740 GTTATCTGACGAGCTACGCTGGCTGGATCTGCTTCCCGCGCGGTATTTGGAG 2799
QY 3150 CTGCTGCTCAAATACCGAGAGGTTCGACGCTGATCTCTCCACCCACCACTGGATGAG 3209
Db 2800 CTGCTGCTCAAATACCGAGAGGTTCGACGCTGATCTCTCCACCCACCACTGGATGAG 2859
QY 3210 GCAGAGTCTGGAGACCTGTGTGGTGTGGTGGCAGGTGGCGCTGTGTGCTGTGGC 3269
Db 2860 GCAGAGCTGTGGAGACCTGTGTGGCCTGTGGTGGCAGGTGGCGCTGTGTGCTGTGGC 2919
QY 3270 TCCCACTCTCTGCGGCTGACCTGGCTCCGGCTACTACTGACGCTGGTGAAGGCC 3329
Db 2920 TCCCACTCTCTGCGGCTGACCTGGCTCCGGCTACTACTGACGCTGGTGAAGGCC 2979
QY 3330 CGCTGCCCCGTACCAATGAGAGGCTGACACTGACATGAGGAGGAGTGTGGACAC 3389
Db 2980 CGCTGCCCCGTACCAATGAGAGGCTGACACTGACATGAGGAGGAGTGTGGACAC 3039
QY 3390 AGCAGGAAAGAAATGACCGACCGGACAGAGTGGCACTCTCTCAGCTGTGGCC 3449
Db 3040 AGCAGGAAAGAAATGACCGACCGGACAGAGTGGCACTCTCTCAGCTGTGGCC 3099
QY 3450 CTGGTACAGACTGGGTGGCCGGGACCGCTGGTGGAGGAGCTGCCACGAGCTGGTG 3509
Db 3100 CTGGTACAGACTGGGTGGCCGGGACCGCTGGTGGAGGAGTGGCCACGAGCTGGTG 3159
QY 3510 CTGGTGTGCCCTACAGGGTGGCCATGACGGAGCTTCCCACTCTTCCGAGAGCTA 3569

Db 3160 CTGGTGTGCTCTACACGGGTGCCATGACGGCAGCTTGCACCACCTCTTCCGAGAGCTA 3219
QY 3570 GACACCGGCTTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCACTGCTGAG 3629
Db 3220 GACACCGGCTTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCACTGCTGAG 3279
QY 3630 GAGAT--CTTCTCTGAAGGTGGTGGAGGAGTGTCTGCGGACACAGATATGGAGGATGGC 3686
Db 3280 GAGGTATCTTCTGAAAGTGTGGAGGAGTGTCTGCGGACACAGATATGGAGGATGGC 3339
QY 3687 AGCTGGGGCAGCACCTATGCAAGGCATTTGCTGGCTAGAGTAACCTCTGCGGCTCAAG 3746
Db 3340 AGCTGGGGCAGCACCTATGCAAGGCATTTGCTGGCTAGAGCTAACCCTTACGGCTCAAG 3399
QY 3747 ATGCCGCCACAGAGACAGGCTTGGAGAACGGGGAACCACTGGGTGAGGCCACAGAGACT 3806
Db 3400 ATGCCGCCACAGAGACAGGCTTGGAGAACGGGGAACCACTGGGTGAGGCCACAGAGACT 3459
QY 3807 GACCAGGGCTTGGGGCCAGACGCGCTGGGCTACAGGGCTTGGGCTGGGCACTGACCCGCCAG 3866
Db 3460 GACCAGGGCTTGGGGCCAGACGCGCTGGGCTTGGGCTTACAGGGCTGGGCACTGACCCGCCAG 3519
QY 3867 CAGCTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGGCCGCGCAGCGCCGCGGCTG 3926
Db 3520 CAGCTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGGCCGCGCAGCGCCGCGGCTG 3579
QY 3927 TTGCCCCAGATCGTGTGCTGCTGCTTCTTGTGGGCTTGGCCCTCTGCTGTTACGCTCATC 3986
Db 3580 TTGCCCCAGATCGTGTGCTGCTGCTTCTTGTGGGCTTGGCCCTCTGCTGTTTACGCTCATC 3639
QY 3987 GTGCTCTCTTTGGGGCACTACCGGCTCTGCGGCTCAGTCCCACTGATGACGCTGTCTCAG 4046
Db 3640 GTGCTCTCTTTGGGGCACTACCGGCTCTGCGGCTCAGTCCCACTGATGACGCTGTCTCAG 3699
QY 4047 GTGCTCTCTTTTTCAGTGAGGACGCCCCAGGGGACCTTGGACGTGCCCGGCTGTCTGAGGGC 4106
Db 3700 GTGCTCTCTTTTTCAGTGAGGAGGCCCCAGGGGACCTTGGACGTGCCCGGCTGTCTGAGGGC 3759
QY 4107 CTGCTCAGGAGGACGAGCTTGGAGGAGCCCCAGTGCAGATAGCTCCCACTGAGCTCCCA 4156
Db 3760 CTGCTCAGGAGGACGAGCTTGGAGGAGCCCCAGTGCAGATAGCTCCCACTGAGCTCCCA 3819
QY 4157 -----CAGGTTCTCGGCACCAAGAGTTCTGCTGTAAGTGGCC 4193
Db 3820 GAGTGTGACAGCCCCATTTGCTGCAGGTTCTCGGCACCAAGAGTTCTGCTGTAAGTGGCC 3879
QY 4194 AAGGTCTTGGCCAGTGGCAACTGGACCCCAAGTCTTCCATGCCAGCTGCCAGTGPAGC 4253
Db 3880 AAGGTCTTGGCCAGTGGCAACTGGACCCCAAGTCTTCCATGCCAGCTGCCAGTGPAGC 3939
QY 4254 CAGCCCGGTGCCCGGCGCTTGTGCCGACTGCCCGGCTGACGCTGGTGGTCCCGCTCCG 4313
Db 3940 CGGCCCGGTGCCCGGCGCTTGTGCCGACTGCCCGGCTGACGCTGGTGGTCCCGCTCCG 3999
QY 4314 CCCCAGCAGTGCACCGGCTCTGGGGAAGTGGTTTCAAGAACTGACAGCGCCGAACTGTCT 4373
Db 4000 CCCCAGCAGTGCACCGGCTCTGGGGAAGTGGTTTCAAGAACTGACAGCGCCGAACTGTCT 4059
QY 4374 GACTTCTTGGTCAAGACCTTACCGCGCTTGTGGCGCTGAGAGCTGAAGAGTGG 4433
Db 4060 GACTTCTTGGTCAAGACCTTACCGCGCTTGTGGCGCTGAGAGCTGAAGAGTGG 4119
QY 4434 GTGAATGAGGTGAGGTACGAGGCTTCTGCTGGGGGGCGGAGACCCAGGCTTGGCTCG 4493
Db 4120 GTGAATGAGGTGAGGTACGAGGCTTCTGCTGGGGGGCGGAGACCCAGGCTTGGCTCG 4179
QY 4494 GGCACAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGCGCTGCTGAGTCCCTGCTGCTGGC 4553
Db 4180 GGCACAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGCGCTGCTGAGTCCCTGCTGCTGGC 4239
QY 4554 GGGGCCCTCGACCGTCTCTGAAAAACCTCAGAGCTGGGCTCAGAGCTGGATGTCTCAG 4613
Db 4240 GGGGCCCTCGACCGTCTCTGAAAAACCTCAGAGCTGGGCTCAGAGCTGGATGTCTCAG 4299

QY 4614 GACAGTCTCAGATCTGGTTTCAACAAACAAAGGCTGCACTCCATGGTGGCCCTTTGTCAAC 4673
DB |||||
QY 4300 GACAGTCTCAGATCTGGTTTCAACAAACAAAGGCTGCACTCCATGGTGGCCCTTTGTCAAC 4359
DB |||||
QY 4674 CGAGCCAGCAACGCAATCTCCGTGCTCACCCTGCGCCCAAGGCGCGCCGCGCCAGCCAC 4733
DB |||||
QY 4360 CGAGCCAGCAACGCAATCTCCGTGCTCACCCTGCGCCCAAGGCGCGCCGCGCCAGCCAC 4419
DB |||||
QY 4734 AGCATCACCACACTCAACACACCCCTTGAACCTCACCAGGAGCAGCTGTCTGAGGCTGCA 4793
DB |||||
QY 4794 CTGATGGGCTCTCCGGTGAAGCTCTCGTCTCCATCTGATGCTGCTTTGGCCATGTCCTTT 4853
DB |||||
QY 4480 CTGATGGGCTCTCCGGTGAAGCTCTCGTCTCCATCTGATGCTGCTTTGGCCATGTCCTTT 4539
DB |||||
QY 4854 GTCCCGGCGAGCTTCACTCTGCTCTCATTTGAGGAGCGAGTCAACCGAGCCCAAGCACCTG 4913
DB |||||
QY 4540 GTCCCGGCGAGCTTCACTCTGCTCTCATTTGAGGAGCGAGTCAACCGAGCCCAAGCACCTG 4599
DB |||||
QY 4914 CAGCTCATGGGGGCTGTCCCCACACCTCTACTGGCTTGGCAACTTTCTCTGGGACATG 4973
DB |||||
QY 4600 CAGCTCATGGGGGCTGTCCCCACACCTCTACTGGCTTGGCAACTTTCTCTGGGACATG 4659
DB |||||
QY 4974 TGTAACTACTTGGTGGCCAGCATGCAATGCTGCTCATCTTTCTGGCCCTTCCAGCAGAG 5033
DB |||||
QY 4660 TGTAACTACTTGGTGGCCAGCATGCAATGCTGCTCATCTTTCTGGCCCTTCCAGCAGAG 4719
DB |||||
QY 5034 GCATATGGGCGCCCTGCCAACTGCTGCTCTCTCTGCTGCTGCTACTGATGCTGCTGG 5093
DB |||||
QY 4720 GCATATGGGCGCCCTGCCAACTGCTGCTCTCTCTGCTGCTGCTACTGATGCTGCTGG 4779
DB |||||
QY 5094 TCGATCACACGCTCATGTACCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5153
DB |||||
QY 4780 TCGATCACACGCTCATGTACCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4839
DB |||||
QY 5154 GTGGTGTCTCAGTCAATAAATCTTTTATTTGGCATCAATGGAAGCATGGCCACCTTTGTG 5213
DB |||||
QY 4840 GTGGTGTCTCAGTCAATAAATCTTTTATTTGGCATCAATGGAAGCATGGCCACCTTTGTG 4899
DB |||||
QY 5214 CTGAGCTCTCTCTGATCAGAGTGTGAGAGGTGAGCGGATCTTGAACAGTCTTC 5273
DB |||||
QY 4900 CTGAGCTCTCTCTGATCAGAGTGTGAGAGGTGAGCGGATCTTGAACAGTCTTC 4959
DB |||||
QY 5274 CTATCTTCCCCACTCTCTGCTTGGGCGGGGCTCATTTGACATGGTGGGAACAGGCCC 5333
DB |||||
QY 4960 CTATCTTCCCCACTCTCTGCTTGGGCGGGGCTCATTTGACATGGTGGGAACAGGCCC 5019
DB |||||
QY 5334 ATGGCTGATGCTTTGAGCGCTTG-----GGAGACAGGCAAGTTCCAG 5375
DB |||||
QY 5020 ATGGCTGATGCTTTGAGCGCTTGCGGTGAGAACTTCTCTGTCAGGAGACAGGCAAGTTCCAG 5079
DB |||||
QY 5376 TCACCCCTGCGCTGGAGAGTGTGGCAAGAACTCTTGGCCATGGTGAATACAGGGGCC 5435
DB |||||
QY 5080 TCACCCCTGCGCTGGAGAGTGTGGCAAGAACTCTTGGCCATGGTGAATACAGGGGCC 5139
DB |||||
QY 5436 CTCTCTCTCTCTCACAATCTGCTGAGCAGCAGGAGCACTCTCTGCGCAGCCAGG 5495
DB |||||
QY 5140 CTCTCTCTCTCTCACAATCTGCTGAGCAGCAGGAGCACTCTCTGCGCAGCCAGG 5199
DB |||||
QY 5496 GTGAGGCTCTCTGCACTCTCTGGAGAGGAGGAGGAGTGTAGCCCTGTAACCGGAGCGG 5555
DB |||||
QY 5200 GTGAGGCTCTCTGCACTCTCTGGAGAGGAGGAGGAGTGTAGCCCTGTAACCGGAGCGG 5259
DB |||||
QY 5556 GTGTCCAAAGAGCACCAGGGGATGTGTGCTGCTGAGAACTTGAACAAAGTATAC 5615
DB |||||
QY 5260 GTGTCCAAAGAGCACCAGGGGATGTGTGCTGCTGAGAACTTGAACAAAGTATAC 5319
DB |||||
QY 5616 CGTGGGAGAGGATGCGAGCTGTGACCGCTGTGCTGGGATTTCCCTCTGTGAGTGT 5675
DB |||||
QY 5320 CGTGGGAGAGGATGCGAGCTGTGACCGCTGTGCTGGGATTTCCCTCTGTGAGTGT 5379
DB |||||

RESULT 11

AX202220

LOCUS

DEFINITION Sequence 3 from Patent WO0153490.

AX202220

5669 bp

DNA

linear

PAT 30-AUG-2001

QY 5676 TTTGGGCTGCTGGGTGTGAATGGAGCAGGG--AAGACGCTCCACGTTTCGATGTTGACGGG 5734
DB |||||
QY 5380 TTTGGGCTGCTGGGTGTGAATGGAGCAGGAAAGACGTCACGTTTCGATGTTGACGGG 5439
DB |||||
QY 5735 GGACACATTTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCA-----CAGC 5780
DB |||||
QY 5440 GGACACATTTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACAGGTGAGGGGTGCAGC 5499
DB |||||
QY 5781 GTGGCCCGGGAAACCCAGTGTCTGCGCACCTCAGCATGGGATCTGCCCTCAATCCGATGCC 5840
DB |||||
QY 5500 GTGGCCCGGGAAACCCAGTGTCTGCGCACCTCAGCATGGGATCTGCCCTCAATCCGATGCC 5559
DB |||||
QY 5841 ATCTTTGAGTGTCTGACGGGCGCGAGCACTCTGGAGCTGTCTGCGGCTGCGCGGTGTC 5900
DB |||||
QY 5560 ATCTTTGAGTGTCTGACGGGCGCGAGCACTCTGGAGCTGTCTGCGGCTGCGCGGTGTC 5619
DB |||||
QY 5901 CCGGAGGCGCCAGTGTGCGGAGCAGCTGCTGAGCTGAGGCTGTGCGGCTGCGGCTCTCATGG 5960
DB |||||
QY 5620 CCGGAGGCGCCAGTGTGCGGAGCAGCTGCTGAGCTGAGGCTGTGCGGCTGCGGCTCTCATGG 5679
DB |||||
QY 5961 TACGACAGACCGGCTTGCAGGCACTTACAGCGGAGGGAACAAACGCAAGCTGGCGACGGCC 6020
DB |||||
QY 5680 TACGACAGACCGGCTTGCAGGCACTTACAGCGGAGGGAACAAACGCAAGCTGGCGACGGCC 5739
DB |||||
QY 6021 CTGGCGCTGTGGGACCCAGCGCTGTGTTTCTGGACGAGCCCAACAGCATATGAC 6080
DB |||||
QY 5740 CTGGCGCTGTGGGACCCAGCGCTGTGTTTCTGGACGAGCCCAACAGCATATGAC 5799
DB |||||
QY 6081 CCCAGCGCGCGGCTTCTTTTGGNAACAGCTTTTGGCGGCTGTGCGGAGGCGGCTTCA 6140
DB |||||
QY 5800 CCCAGCGCGCGGCTTCTTTTGGNAACAGCTTTTGGCGGCTGTGCGGAGGCGGCTTCA 5859
DB |||||
QY 6141 GTGATGCTCACTTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATC 6200
DB |||||
QY 5860 GTGATGCTCACTTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATC 5919
DB |||||
QY 6201 ATGTGTAATGGCGGTTCCGCTGCTGGGACCGCAACATCTCAAGGCGAGATTCGCG 6260
DB |||||
QY 5920 ATGTGTAATGGCGGTTCCGCTGCTGGGACCGCAACATCTCAAGGCG----- 5970
DB |||||
QY 6261 GCGGCTCACACTGACCTCGGGTGCCTGCGGCGCAGGTCAGGTCAGGCGGCGGCTTC 6320
DB |||||
QY 5971 -----AGGTCCCAGCGGCGGCGCTTC 5994
DB |||||
QY 6321 GTGGCGCGGAGTTCCTTGGTGGAGCTGCGGAGGACATGAGGCGGCTGCGCTTC 6380
DB |||||
QY 5995 GTGGCGCGGAGTTCCTTGGGCGGAGCTGCGGAGGACATGAGGCGGCTGCGCTTC 6054
DB |||||
QY 6381 CAGTGGCGCGGAGGCGCTGCGGCTGCGGCGGCTTTTGGAGAGCTGCGGCTGCAC 6440
DB |||||
QY 6055 CAGTGGCGCGGAGGCGCTGCGGCTGCGGCGGCTTTTGGAGAGCTGCGGCTGCAC 6114
DB |||||
QY 6441 GGCGCAGAGCAGGCGTGGAGGCTTTTCCGTGAGCAGACATGCTGAGGAGGATTC 6500
DB |||||
QY 6115 GGCGCAGAGCAGGCGTGGAGGCTTTTCCGTGAGCAGACATGCTGAGGAGGATTC 6174
DB |||||
QY 6501 TTGTACTTCTCCAAAGGACAGGGAAGGACCGAGGACCCGAGGAGCAGAGGAGCAGGA 6560
DB |||||
QY 6175 TTGTACTTCTCCAAAGGACAGGGAAGGACCGAGGACCCGAGGAGCAGAGGAGCAGGA 6234
DB |||||
QY 6561 GTGGAGTGGACCCCGGCGGAGGCTGCGAGCACCACCAACCGCTCAGGCTTCCTCGAT 6620
DB |||||
QY 6235 GTGGAGTGGACCCCGGCGGAGGCTGCGAGCACCACCAACCGCTCAGGCTTCCTCGAT 6294
DB |||||
QY 6621 GACCTTAGCACTGCGGAGACTGTGCTTGA 6650
DB |||||
QY 6295 GACCTTAGCACTGCGGAGACTGTGCTTGA 6324
DB |||||

QY 2070 CCTACAGCCACCCGGGCGTGTCTTCTGTTCTTGGAGCGCTTCGCGGTGGCCACCGGTG 2129
DB |||||
1861 CCTCTGAGCCACCCGGGCGTGTCTTCTGTTCTTGGAGCGCTTCGCGGTGGCCACCGGTG 1920
QY 2130 ACCCAGAGCTTCCCTGCTCAGCGCCTTCTTCCCGGCGCCACCTGGCTGGCGCTTCGCGC 2189
DB |||||
1921 ACCCAGAGCTTCCCTGCTCAGCGCCTTCTTCTCCCGCGCCAACTGGCTGGCGCCTTCGCGC 1980
QY 2190 GGCCTGGGCTACTTCTCCCTCTACCTGGCCCTACGTGCTGTGTGTGGCTTGGCGGACCGG 2249
DB |||||
1981 GGCCTGGGCTACTTCTCCCTCTACCTGGCCCTACGTGCTGTGTGTGGCTTGGCGGACCGG 2040
QY 2250 CTGCCCGCGGTGGCGGCTGGCCGCGAGCCTGCTGTGCGCGGTGGCCTTCGGCTTCGGC 2309
DB |||||
2041 CTGCCCGCGGTGGCGGCTGGCCGCGAGCCTGCTGTGCGCGGTGGCCTTCGGCTTCGGC 2100
QY 2310 TGGAGAGCCTGGCTCTGCTGGAGAGCAGGGGCGAGGGGCGGCGAGTGCGACACAGTGGC 2369
DB |||||
2101 TGGAGAGCCTGGCTCTGCTGGAGAGCAGGGGCGAGGGGCGGCGAGTGCGACACAGTGGG 2160
QY 2370 ACCGGGCTACGGCAGAGCTTCTAGCCTGGCCCAAGGTCTCTGGGCTTCTGTGCTGGAC 2429
DB |||||
2161 ACCGGGCTACGGCAGAGCTTCTAGCCTGGCCCAAGGTCTCTGGGCTTCTGTGCTGGAC 2220
QY 2430 GCGGCGCTTACGGCCTCGCCACTCTGTGTAAGCTGTGTGCGCCAGGCGAGTACGGG 2489
DB |||||
2221 GCGGCGCTTACGGCCTCGCCACTCTGTGTAAGCTGTGTGCGCCAGGCGAGTACGGG 2280
QY 2490 ATCCCTGAACATGGAATTTCTTTTCGGAGGAGCTACTGGTGGGAGCTTCGGGCCCCC 2549
DB |||||
2281 ATCCCTGAACATGGAATTTCTTTTCGGAGGAGCTACTGGTGGGAGCTTCGGGCCCCC 2340
QY 2550 AAGAGTCAGGCCCCCTTGCCCAACCCCGCTGGAACCAAGGTGCTGTGTAAGAGGACCG 2609
DB |||||
2341 AAGAGTCAGGCCCCCTTGCCCAACCCCGCTGGAACCAAGGTGCTGTGTAAGAGGACCG 2400
QY 2610 CCGGCTGAGTCTCGGCTTTCGCTTTCGAGCCTGGAGAGCGCTTTCCTGGAGCGCG 2669
DB |||||
2401 CCGGCTGAGTCTCGGCTTTCGCTTTCGAGCCTGGAGAGCGCTTTCCTGGAGCGCG 2460
QY 2670 CAGCAGCCCTCGGGGCTCAGCTGAGCTTCTACAGGGGCGCATCAGCGCTTCTG 2729
DB |||||
2461 CAGCAGCCCTCGGGGCTCAGCTGAGCTTCTACAGGGGCGCATCAGCGCTTCTG 2520
QY 2730 GGCCAAACGGGGCGGCAAGACCAACCTGTCTTGTGAGTGGCTCTTCCACCC 2789
DB |||||
2521 GGCCAAACGGGGCGGCAAGACCAACCTGTCTTGTGAGTGGCTCTTCCACCC 2580
QY 2790 AGTGTGGCTCTGCTTTCATCTCGGGCGACAGCTTCGCTCAGCATGGCGCCATCCGG 2849
DB |||||
2581 AGTGTGGCTCTGCTTTCATCTCGGGCGACAGCTTCGCTCAGCATGGCGCCATCCGG 2640
QY 2850 CCCACCTGGGCGCTGTCTCAGTACAAAGTGTGTTGACATGCTGACCGTGGACGAG 2909
DB |||||
2641 CCCACCTGGGCGCTGTCTCAGTACAAAGTGTGTTGACATGCTGACCGTGGACGAG 2700
QY 2910 CAGCTGTGTTCTATGGGCGCTGAAGGCTCTGAGTGGCGCTGTAGTGGGCGCCGAGCAG 2969
DB |||||
2701 CAGCTGTGTTCTATGGGCGCTGAAGGCTCTGAGTGGCGCTGTAGTGGGCGCCGAGCAG 2760
QY 2970 GACCGTCTGCTGAGGATGTGGGCTGGTCTCCAAGCAGAGTGTGCGACCTCGCCACCTC 3029
DB |||||
2761 GACCGTCTGCTGAGGATGTGGGCTGGTCTCCAAGCAGAGTGTGCGACCTCGCCACCTC 2820
QY 3030 TCTGTGGATGCAACGGAAGCTGCTGGGCGCATTTGCTTGTGGGCGGCTCCCAAGTT 3089
DB |||||
2821 TCTGTGGATGCAACGGAAGCTGCTGGGCGCATTTGCTTGTGGGCGGCTCCCAAGTT 2880
QY 3090 GTTATCTGGAGAGCCTACGGCTGGGCTGGATCTCTGCTCCCGCGGCTATTTGGAG 3149
DB |||||
2881 GTTATCTGGAGAGCCTACGGCTGGGCTGGATCTCTGCTCCCGCGGCTATTTGGAG 2940

QY 3150 CTGCTGTCAAATAACCGAGAAAGGTGCGACGCTGATCTCTCTCCACCCACCTGGATGAG 3209
DB |||||
2941 CTGCTGTCAAATAACCGAGAAAGGTGCGACGCTGATCTCTCTCCACCCACCTGGATGAG 3000
QY 3210 GCAGAGCTGTGGGAGACCGGTGTGGCTGTGGTGGGAGGTGGCGCTGTGCTGTGGC 3269
DB |||||
3001 GCAGAGCTGTGGGAGACCGGTGTGGCTGTGGTGGGAGGTGGCGCTGTGCTGTGGC 3060
QY 3270 TCCCACTCTTCTTCCGCGCGCTCACCTGGGCTCCGGCTACTACTCTGACGCTGGTGAAGGCC 3329
DB |||||
3061 TCCCACTCTTCTTCCGCGCGCTCACCTGGGCTCCGGCTACTACTCTGACGCTGGTGAAGGCC 3120
QY 3330 CGCTGTCCCTTGACCAACATGAGAAGCTGACACTGATGAGGGGAGTGTGACACC 3389
DB |||||
3121 CGCTGTCCCTTGACCAACATGAGAAGCTGACACTGATGAGGGGAGTGTGACACC 3180
QY 3390 AGGAGGAAAGAAAGAAATGGCAGCGGCGAGAGGTGGGAGTCTCCTCAGCTGTGGCC 3449
DB |||||
3181 AGGAGGAAAGAAAGAAATGGCAGCGGCGAGAGTGGGAGTCTCCTCAGCTGTGGCC 3240
QY 3450 CTGGTACAGCACTGGGTGCCCGGGCGACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTG 3509
DB |||||
3241 CTGGTACAGCACTGGGTGCCCGGGCGACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTG 3300
QY 3510 CTGTGTGCTTACACGCGGTGCCCATGACGCGAGCTTTCGCCACACTCTTCGAGAGCTA 3569
DB |||||
3301 CTGTGTGCTTACACGCGGTGCCCATGACGCGAGCTTTCGCCACACTCTTCGAGAGCTA 3360
QY 3570 GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATTCGACACACGAGCTCGAG 3629
DB |||||
3361 GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATTCGACACACGAGCTCGAG 3420
QY 3630 GAGATCTTCTTGAAGGTGGTGGAGGTGTCTTCGGGACACAGATATGGAGGATGGCAGC 3689
DB |||||
3421 GAGATCTTCTTGAAGGTGGTGGAGGTGTCTTCGGGACACAGATATGGAGGATGGCAGC 3480
QY 3690 TGGCGGCGAGCACTTATGACAGGCACTTGTGGCTTAGAGCTAACCTTCGGCTCAAGATG 3749
DB |||||
3481 TGGCGGCGAGCACTTATGACAGGCACTTGTGGCTTAGAGCTAACCTTCGGCTCAAGATG 3540
QY 3750 CCGCCACAGGAGACAGCGCTGGAGAACGGGAAACGAGTGGGTGAGCCCCAGAGCTGAC 3809
DB |||||
3541 CCGCCACAGGAGACAGCGCTGGAGAACGGGAAACGAGTGGGTGAGCCCCAGAGCTGAC 3600
QY 3810 CAGGCTCTGGGCGAGACGCGTGGCGGCTTACAGGCTGGGCACTCACCCGCGAGCAG 3869
DB |||||
3601 CAGGCTCTGGGCGAGACGCGTGGCGGCTTACAGGCTGGGCACTCACCCGCGAGCAG 3660
QY 3870 CTCAGGCGCTTCTTCAAGCGCTTCTGTGCTTCCCGCGAGCGCGCGGCGCTGTTC 3929
DB |||||
3661 CTCAGGCGCTTCTTCAAGCGCTTCTGTGCTTCCCGCGAGCGCGCGGCGCTGTTC 3720
QY 3930 GCGCAGATCTGCTGCTGCGCTTCTTGTGGGCGCTTGGCCCTCGTGTTCAGGCTCATCTG 3989
DB |||||
3721 GCGCAGATCTGCTGCTGCGCTTCTTGTGGGCGCTTGGCCCTCGTGTTCAGGCTCATCTG 3780
QY 3990 CCTCTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGTACGCTGCTCAGGTG 4049
DB |||||
3781 CCTCTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGTACGCTGCTCAGGTG 3840
QY 4050 TCTTCTTTCAGTAGGAGCGCCCGAGGAGCCCTTGGAGCTTCCCGGCTGCTCGAGCGCTG 4109
DB |||||
3841 TCTTCTTTCAGTAGGAGCGCCCGAGGAGCCCTTGGAGCTTCCCGGCTGCTCGAGCGCTG 3900
QY 4110 CTGACAGGCGAGCTGAGGAGCGCCCGAGTGGAGCTAGCTCCACAGGTTCTTCGGCA 4169
DB |||||
3901 CTGACAGGCGAGCTGAGGAGCGCCCGAGTGGAGCTAGCTCCACAGGTTCTTCGGCA 3960
QY 4170 CCAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGAGACCCAGAGTCT 4229
DB |||||
3961 CCAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGAGACCCAGAGTCT 4020
QY 4230 CCATCCCGAGCCTGCGAGTGTAGCCAGCGCGGCTGCGGCGCTGCTGCCCGAGCTGCCG 4289

4021	Db	CCATCCCCAGCCTGCCAGTGTAGCCAGCCGGGTGCCCGGCCCTGCTGCCCCAGCTGCCG	4080
4290	Qy	GCTGCAGCTGGTGGTCCCTCCGCCCCAGGCAGTGACGGCTCTGGGGAAGTGGTTCAG	4349
4081	Db	GCTGCAGCTGGTGGTCCCTCCGCCCCAGGCAGTGACCGGCTCTGGGGAAGTGGTTCAG	4140
4350	Qy	AACCTGACAGCCGCGAACCTGTCTGACTTCCCTGGTCAAGACTTACC CGCCCTGGTGC GC	4409
4141	Db	AACTGACAGCCGCGAACCTGTCTGACTTCTGGTCAAGACTTACC CGCCCTGGTGC GC	4200
4410	Qy	CAGGGCTGAAGACTTAAGAAAGTGGTGAATGAGGTACAGGTACGGAGGCTTCTCGCTGGG	4469
4201	Db	CAGGGCTGAAGACTTAAGAAAGTGGTGAATGAGGTACAGGTACGGAGGCTTCTCGCTGGG	4260
4470	Qy	GGCGAGACCCAGGCCCTGCCCTCGGGCCAAGAGTTGGGCCGCTCAGTGTGAGGAGTTGTGG	4529
4261	Db	GGCGAGACCCAGGCCCTGCCCTCGGGCCAAGAGTTGGGCCGCTCAGTGTGAGGAGTTGTGG	4320
4530	Qy	GGCGTGTGAGTCCCTCGCTGGCGGGGCCCTCGACCGTCTCTCTGAAAAACCTCACAGCC	4589
4321	Db	GGCGTGTGAGTCCCTCGCTGGCGGGGCCCTCGACCGTCTCTCTGAAAAACCTCACAGCC	4380
4590	Qy	TGGGCTCACAGCCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAAGGCTGG	4649
4381	Db	TGGGCTCACAGCCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAAGGCTGG	4440
4650	Qy	CACTCCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCCGTGCTCACTGCCGCC	4709
4441	Db	CACTCCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCCGTGCTCACTGCCGCC	4500
4710	Qy	CCAGGCCGGCCCGCCACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACACC	4769
4501	Db	CCAGGCCGGCCCGCCACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACACC	4560
4770	Qy	AAGGAGCAGCTGTCTGAGGCTGCACCTGATGGCCCTCCTCGGTGGAAGCTCTCGTCTCCATC	4829
4561	Db	AAGGAGCAGCTGTCTGAGGCTGCACCTGATGGCCCTCCTCGGTGGAAGCTCTCGTCTCCATC	4620
4830	Qy	TGTGTGCTCTTTGGCATGTCTTTTGTCCGGCCAGCTTCACTCTTGTGCTCATTTAGGAG	4889
4621	Db	TGTGTGCTCTTTGGCATGTCTTTTGTCCGGCCAGCTTCACTCTTGTGCTCATTTAGGAG	4680
4890	Qy	CGAGTCAACCGAGCCAGACCTGCAGCTCATGGGGGCCCTGTCCGCCACCTCTACTGG	4949
4681	Db	CGAGTCAACCGAGCCAGACCTGCAGCTCATGGGGGCCCTGTCCGCCACCTCTACTGG	4740
4950	Qy	CTTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCAGCATGCATCGTGGTGCTC	5009
4741	Db	CTTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCAGCATGCATCGTGGTGCTC	4800
5010	Qy	ATCTTTCTGGCTTCCAGCAGAGGGCATATGTGGCCCTCCGCAACCTGCCTGCTCTCCTG	5069
4801	Db	ATCTTTCTGGCTTCCAGCAGAGGGCATATGTGGCCCTCCGCAACCTGCCTGCTCTCCTG	4860
5070	Qy	CTGTTGCTACTACTGTATGCTGGTCGATCACACCGCTCATGTACCCAGCGCTCTCTTTC	5129
4861	Db	CTGTTGCTACTACTGTATGCTGGTCGATCACACCGCTCATGTACCCAGCGCTCTCTTTC	4920
5130	Qy	TTCTCCGTCGCCAGCACAGGCTATGTGGTGCTCACCTGCATAAACCTCTTTATTGGCATC	5189
4921	Db	TTCTCCGTCGCCAGCACAGGCTATGTGGTGCTCACCTGCATAAACCTCTTTATTGGCATC	4980
5190	Qy	AATGGAAGCATGGCCACTTTTGTGCTTGAGCTTCTCTGATCAGAAAGCTGCAGGAGGTG	5249
4981	Db	AATGGAAGCATGGCCACTTTTGTGCTTGAGCTTCTCTGATCAGAAAGCTGCAGGAGGTG	5040
5250	Qy	AGCCGGATCTTGAACAGGCTCTCTTATCTTCCGCCACTTCTGCTTGGGCCCGGGGCTC	5309
5041	Db	AGCCGGATCTTGAACAGGCTCTCTTATCTTCCGCCACTTCTGCTTGGGCCCGGGGCTC	5100
5310	Qy	ATTGACATGGTGGCGAACCAGGCCATGGCTGATGCTCTTGAGCGCTTGGGAGACAGGCAG	5369

Db	5101	ATTGACATGGTGGCGAACGAGGCCATGGCTGATGCTCTTTGAGCGCTTGGAGACAGGCAG	5161
Qy	5370	TTCCAGTCAACCCCTCGCGCTGGAGGTGGTCGGCAAGAACTCTTTGGGCCATGGTGATACAG	5429
Db	5161	TTCCAGTCAACCCCTCGCGCTGGAGGTGGTCGGCAAGAACTCTTTGGGCCATGGTGATACAG	5220
Qy	5430	GGGCCCCCTTCTCTTCTTTCACACTACTGCTGCAGCACCGCAAGCCAACTCTCTGCCACAG	5489
Db	5221	GGGCCCCCTTCTCTTCTTTCACACTACTGCTGCAGCACCGCAAGCCAACTCTCTGCCACAG	5280
Qy	5490	CCGAGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGGACGAGGATGTAGGCCCTGTAACGG	5549
Db	5281	CCGAGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGGACGAGGATGTAGGCCCTGTAACGG	5340
Qy	5550	GAGCGGTGTCGCAAGGAGCCACCGAGGGGATGCTGTTGCTGTGAGGAACTTGACCAAG	5609
Db	5341	GAGCGGTGTCGCAAGGAGCCACCGAGGGGATGCTGTTGCTGTGAGGAACTTGACCAAG	5400
Qy	5610	GTATACCGTGGGAGAGGATGCCAGCTGTTTGACCGCTTGTGCTCTGGGGATTTCCCCCTGGT	5669
Db	5401	GTATACCGTGGGAGAGGATGCCAGCTGTTTGACCGCTTGTGCTCTGGGGATTTCCCCCTGGT	5460
Qy	5670	GAG 5672	
Db	5461	GAG 5463	
RESULT 12			
AX644625	AX644625	4653 bp	DNA
LOCUS	Sequence 7 from Patent WO0209108.	linear	PAT 27-FEB-2003
DEFINITION	AX644625		
ACCESSION	AX644625		
VERSION	AX644625.1	GI:28610639	
KEYWORDS			
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Shutter,J. and Ulias,L.		
TITLE	Alp-binding cassette transporter-like molecules and uses thereof		
JOURNAL	Patent: WO 0209108-A 7 12-DEC-2002;		
FEATURES	Angen, Inc. (US)		
source	Location/Qualifiers		
	1..4653		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	1..4653		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CA068133.1"		
	/db_xref="GI:28610640"		
	/translation="MVCLGFGQSGAPLVSVQNHCPGCLSPQBSLGLALGQAQEPHLSH LLEAAGDLAQELASRLPDLRALLQRPCTGSPLELLSEALCSVRGSPSTVGPSLNNW YEASDLMELVQBPESALPDSLSAPCSLEIGALDSHPLSRLLRWRLLKPLILGKLPA PDPTRFKMAQVNRFEELTLRDVREWEMLGRIPTFMNDSNNVAMQLRLQLQMDQ LILWLGSMFLSLCGLFLLSAALLVLKGLDILPVSHGVVFLFLAFAVATVTSQFLLS EGRRPPRCGRDHREALRFLDPGSGGYSQDDAHVGLVGTGRVTECLSLDLKLEA APSEAAVSRALQLLAHRRFWAGVFLGPDSDPTEHPTDILGPHRVTKIRINDVID VTRNKRIDRFPDGPAADELTLDRVYWGGFVQLQDLVERAAVRVLSGANPRAGLYLO QMPYPCVDVDFLRVLRSLLPLFLTAWIVSVTLTKAVVRKEKRLTRDTRMAGLSR AVLWLGSMFLSLCGLFLLSAALLVLKGLDILPVSHGVVFLFLAFAVATVTSQFLLS SAFESRANLAAACGGLAYFSLYPYVLVAVWRDLRPAAGRVAAASLLSPVAFGCESL ALLIEQEGQAHWNVGRTPADRVFSLAQVSGILLDDAALYGLATWLEAVCRGQYGP EPNWFRRRSYWCQPPPKSPACPTPDKPVLVBEAPGLSPGVSFVSRSLEKFPGPS QPALRGLSLDFYQGHITATFLHNGAGTKTTLISGLFPPSGSAVTLGHDDVRSMAA IRHLPGVCPQYNVLFDMLTVDHEWVYFGRKLGSAAVGPEDRLQDVLGKQSVQ TRHLSGQMOKLSVAVAFVGGSSWILDEPTAGVDPASRRGIVWELLKYREGRTLILS THLDEALLGDVAVVAGGRICCGSPFLRRHLSGYYVTLVKARPLVTTEKADT GMSVDVTRQKNGSGSRVCTPQALLVQHWVPGARLVEELPHELVLVLPYTGADH DSFATFREUDTLAEURLTGYISDTSLEEIFLKVVEECAADTDMEDGCGQHLCGTG IAGDLVTRLRKMPQTALENPGFAGSAPETDQSGPDAVGRVQGWALTROQLQALL KRFLVTRSRRGIFAQITVLPALFVGLVFSLIVPPFGHYPALELSPTMYGAQVFFS		

EDAPDPCBARLLEALLQEBGLEBPPVOHSHRFSAPBPVPAEVAKVLASGNWTPESP
FACQSCQBARLLEALLQEBGLEBPPVOHSHRFSAPBPVPAEVAKVLASGNWTPESP
OGLKTKRWNVRRVYGGFSLGRRDPLGSLGRSVEBELWALLSPLPGCALDRVLKNI
TAWHSLDAQSLSKITFFNNKQWHSMVAFVNRASNAILRAHLPPGPBARHAHSITTLNHP
LNLTKEOLSEALMASSVDVLVSI CVVPMFSVPASFTLVLI EBRVTRAKHLQMLGGL
SPTLYWLGFLWDMCNLYLPACIIVLIEFLAQORAYVAPANILPALLILLLLLYGRQAVP
VTPALGGRQBPGLHGDGTGAPLPLSLHTTAAAPKPTPATPAQGEVSATPGRGRCSP"

ORIGIN

Query Match		64.4%;	Score	4381.6;	DB	6;	Length	4653;			
Best Local Similarity		99.9%;	Pred.	No.	0;						
Matches		4384;	Conservative	0;	Mismatches	4;	Indels	0;			
		Gaps		0;							
QY	702	CGCAGCGAATCCCTGGGTTGGCACTGGGCCAAGCCAGAGCCCTTGACAGCTTGTG	761								
DB	79	CCCAGGAAATCCCTGGGTTGGCACTGGGCCAAGCCAGAGCCCTTGACAGCTTGTG	138								
QY	762	GAGGCCGTGAGAACCTGGCCCGAGGAGCTCCTGGCGCTGCGACCTGGTGGAGCTTCGG	821								
DB	139	GAGGCCGTGGGACCTGGCCCGAGGAGCTCCTGGCGCTGCGACCTGGTGGAGCTTCGG	198								
QY	822	GCACCTGTCAGAGACCCCGAGGAGACAGCGGGCCCTCTGAGTTGCTGTGAGAGGCCCTC	881								
DB	199	GCACCTGTCAGAGACCCCGAGGAGACAGCGGGCCCTCTGAGTTGCTGTGAGAGGCCCTC	258								
QY	882	TGCAGTGTACGGGACCTAGCAGCACAGTGGGGCCCTCTCCCTCAACTGGTACGAGGCTAGT	941								
DB	259	TGCAGTGTACGGGACCTAGCAGCACAGTGGGGCCCTCTCCCTCAACTGGTACGAGGCTAGT	318								
QY	942	GACCTGATGGAGCTGGTGGGCGAGGACCAAGAAATCCGCCCTGCGCAGACAGCACTGAGC	1001								
DB	319	GACCTGATGGAGCTGGTGGGCGAGGACCAAGAAATCCGCCCTGCGCAGACAGCACTGAGC	378								
QY	1002	CCCGCCTGCTCGAGCTGATTGGAGCCCTGGACAGCCACCCGCTGTCCCGCTGCTCTGG	1061								
DB	379	CCCGCCTGCTCGAGCTGATTGGAGCCCTGGACAGCCACCCGCTGTCCCGCTGCTCTGG	438								
QY	1062	AGACGCTCAAGCCTCTGATCTCTCGGAAGCTACTCTTTGACACAGATACACCTTTTACC	1121								
DB	439	AGACGCTCAAGCCTCTGATCTCTCGGAAGCTACTCTTTGACACAGATACACCTTTTACC	498								
QY	1122	CGGAAGCTCATGCCCCAGGTGAACCGGACCTTCGAGGAGCTCACCTCTGAGGGATGTC	1181								
DB	499	CGGAAGCTCATGCCCCAGGTGAACCGGACCTTCGAGGAGCTCACCTCTGAGGGATGTC	558								
QY	1182	CGGAGGTGTGGAGATGCTGGACCCCGGATCTTACCTTCATGAACGACAGTTCCAAT	1241								
DB	559	CGGAGGTGTGGAGATGCTGGACCCCGGATCTTACCTTCATGAACGACAGTTCCAAT	618								
QY	1242	GTGGCCATGCTGCAGCGGCTCTTCAGATGCAAGATGAAGGAAGAGGACCCAGACCT	1301								
DB	619	GTGGCCATGCTGCAGCGGCTCTTCAGATGCAAGATGAAGGAAGAGGACCCAGACCT	678								
QY	1302	GGAGCCCGGACCAATGGAGGCCCTTCGATCTCTTTCTGGACCTCGGAGCGGTGGCTAC	1361								
DB	679	GGAGCCCGGACCAATGGAGGCCCTTCGATCTCTTTCTGGACCTCGGAGCGGTGGCTAC	738								
QY	1362	AGCTGGCAGGACGACACGCTGATGTGGGGACCTGTGTGGGACGCTGGGCCGAGTGACG	1421								
DB	739	AGCTGGCAGGACGACACGCTGATGTGGGGACCTGTGTGGGACGCTGGGCCGAGTGACG	798								
QY	1422	GAGTGCCTGTCTTGGACAAGCTGGAGGGCGGACCCCTCAGAGCAGCCCTGGTGTGCGGG	1481								
DB	799	GAGTGCCTGTCTTGGACAAGCTGGAGGGCGGACCCCTCAGAGCAGCCCTGGTGTGCGGG	858								
QY	1482	GCCCTGCAACTGCTCGCGGAACATCGATTTCTGGGCGGGCGTCTCTTTTGGGACCTGAG	1541								
DB	859	GCCCTGCAACTGCTCGCGGAACATCGATTTCTGGGCGGGCGTCTCTTTTGGGACCTGAG	918								
QY	1542	GACTCTTCAGACCCACAGAGCACCCAAACCCACAGACTGGGGCCCCGGGCACTGGGCTC	1601								
DB	919	GACTCTTCAGACCCACAGAGCACCCAAACCCACAGACTGGGGCCCCGGGCACTGGGCTC	978								

Qy	1602	AAAATCCCATGGACATTTGACGTGTACAGAGGACCAATAAGATCAGGGACAGTTTGG	1661					
Db	979	AAAATCCCATGGACATTTGACGTGTACAGAGGACCAATAAGATCAGGGACAGTTTGG	1038					
Qy	1662	GACCTGGCCAGCGCGGACCCCTGACCGACTGCGCTACGTTGGGGCGGCTTCGTG	1721					
Db	1039	GACCTGGCCAGCGCGGACCCCTGACCGACTGCGCTACGTTGGGGCGGCTTCGTG	1098					
Qy	1722	TACCTGCAAGACCTGGTGGAGCGTGCAGCCGTCGCGCTGCTCAGCGGGCCCAACCCCGG	1781					
Db	1099	TACCTGCAAGACCTGGTGGAGCGTGCAGCCGTCGCGCTGCTCAGCGGGCCCAACCCCGG	1158					
Qy	1782	GCGGCTCTTACTGTCAGCAGATGCCCTTATCCGTGCTATGTGGACGACGTGTCTCGCT	1841					
Db	1159	GCGGCTCTTACTGTCAGCAGATGCCCTTATCCGTGCTATGTGGACGACGTGTCTCGCT	1218					
Qy	1842	GTGCTGAGCCGCTGCTGCGCTCTTCTGAGCGTGGCTGGATCTACTCTCGTGACACTG	1901					
Db	1219	GTGCTGAGCCGCTGCTGCGCTCTTCTGAGCGTGGCTGGATCTACTCTCGTGACACTG	1278					
Qy	1902	ACAGTGAAGGCGGTGGTGGGAGAAAGAGACGCGGCTGCGGACACCATGCGCGCCATG	1961					
Db	1279	ACAGTGAAGGCGGTGGTGGGAGAAAGAGACGCGGCTGCGGACACCATGCGCGCCATG	1338					
Qy	1962	GGGCTCAGCCCGCGGTGCTCTGCTAGGCTAGGCTGCTTCTCAGCTGCGGCTTCCTG	2021					
Db	1339	GGGCTCAGCCCGCGGTGCTCTGCTAGGCTAGGCTGCTTCTCAGCTGCGGCTTCCTG	1398					
Qy	2022	CTCAGCGCGCGCTGCTGCTTCTGCTCAAGCTGGGGACATCTCTCCCTACAGCCAC	2081					
Db	1399	CTCAGCGCGCGCTGCTGCTTCTGCTCAAGCTGGGGACATCTCTCCCTACAGCCAC	1458					
Qy	2082	CCGGCGGTGGTCTTCTGCTTCTGCGAGCTTCTGCGGTGGCCACCGTGCACCCAGAGCTTC	2141					
Db	1459	CCGGCGGTGGTCTTCTGCTTCTGCGAGCTTCTGCGGTGGCCACCGTGCACCCAGAGCTTC	1518					
Qy	2142	CTGCTCAGCGCTTCTTCTCCCGGCCAACTGCTGCGGCTGCGGCGGCTGCGGCTTAC	2201					
Db	1519	CTGCTCAGCGCTTCTTCTCCCGGCCAACTGCTGCGGCTGCGGCGGCTGCGGCTTAC	1578					
Qy	2202	TTCTCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2261					
Db	1579	TTCTCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1638					
Qy	2262	GGCGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2321					
Db	1639	GGCGCGGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1698					
Qy	2322	GCTCTGCTGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACG	2381					
Db	1699	GCTCTGCTGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACG	1758					
Qy	2382	GCAGAGCTTTCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAC	2441					
Db	1759	GCAGAGCTTTCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAC	1818					
Qy	2442	GGCTCTGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2501					
Db	1819	GGCTCTGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1878					
Qy	2502	TGGAATTTTCTTTTCGAGGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2561					
Db	1879	TGGAATTTTCTTTTCGAGGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1938					
Qy	2562	CCTTGCCCGACCCCGCTGAGACCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2621					
Db	1939	CCTTGCCCGACCCCGCTGAGACCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1998					
Qy	2622	CCTGCGGTATCCGTTGCGAGCTGAGAGCGCTTCTCTGGAAGCGCGGCGGCGGCTG	2681					
Db	1999	CCTGCGGTATCCGTTGCGAGCTGAGAGCGCTTCTCTGGAAGCGCGGCGGCGGCTG	2058					
Qy	2682	CGGGGGCTGAGCTGAGCTTCTTACAGGGGCCACATACCGCGCTTCTCTGGGCGGCGG	2741					

Db	2059	CGGGGCTCAGCCTGGACTTCTACCAAGGCAATCACCGCCCTTCTTGGGCCCAACAGGG	2118
Qy	2742	GCCGGCAAGACCAACCCCTGTCTCATCTTGAAGTGGCTCTTCCACCCAGTGTGGCTCT	2801
Db	2119	GCCGGCAAGACCAACCCCTGTCTCATCTTGAAGTGGCTCTTCCACCCAGTGTGGCTCT	2178
Qy	2802	GCCTTCATCTGGGGCCAGCAGCTCCGCTCCAGATGGCCGCAATCCGGCCCACTCTGGGC	2861
Db	2179	GCCTTCATCTGGGGCCAGCAGCTCCGCTCCAGATGGCCGCAATCCGGCCCACTCTGGGC	2238
Qy	2862	GTCTGTCTCAGTACACGCTGTCTTGTGACATGTGACCCGTGGACGACGCTGTGGTTC	2921
Db	2239	GTCTGTCTCAGTACACGCTGTCTTGTGACATGTGACCCGTGGACGACGCTGTGGTTC	2298
Qy	2922	TATGGGGCGGTGAAGGGTGTGAGTGGCGCTGTGATGGGCCCCGAGACGACCGTGTGCTG	2981
Db	2299	TATGGGGCGGTGAAGGGTGTGAGTGGCGCTGTGATGGGCCCCGAGACGACCGTGTGCTG	2358
Qy	2982	CAGGATGTGGGGTGTCTCCAGCAGAGTGTGCAGACTCGCACCTCTCTGGTGGGATG	3041
Db	2359	CAGGATGTGGGGTGTCTCCAGCAGAGTGTGCAGACTCGCACCTCTCTGGTGGGATG	2418
Qy	3042	CAACGGAAGCTGTCCGTGGCCATTGCTTTGTGGGGGCTCCCAAGTTGTTATCTCTGGAC	3101
Db	2419	CAACGGAAGCTGTCCGTGGCCATTGCTTTGTGGGGGCTCCCAAGTTGTTATCTCTGGAC	2478
Qy	3102	GAGCTACGGCTGGGTGGATCTCTGTCTCCCGCGGGATTTTGGAGCTGTCTGCTCAAA	3161
Db	2479	GAGCTACGGCTGGGTGGATCTCTGTCTCCCGCGGGATTTTGGAGCTGTCTGCTCAAA	2538
Qy	3162	TACCGAGAGGTGCAGCTGATCTCTCCACCCACACCTGATGAGGACGAGCTGCTG	3221
Db	2539	TACCGAGAGGTGCAGCTGATCTCTCCACCCACACCTGATGAGGACGAGCTGCTG	2598
Qy	3222	GGAGACCGTGTGGTGTGGCAGGTGGCGCTTGTGTGCTGTGGCTCCCACTCTTC	3281
Db	2599	GGAGACCGTGTGGTGTGGCAGGTGGCGCTTGTGTGCTGTGGCTCCCACTCTTC	2658
Qy	3282	CTGGCGGCTCACTGGGCTCCGGCTACTACTGACGCTGTGTGAAGGCGCGCTGCCCTG	3341
Db	2659	CTGGCGGCTCACTGGGCTCCGGCTACTACTGACGCTGTGTGAAGGCGCGCTGCCCTG	2718
Qy	3342	ACCACCAATGAGAGGCTGACCTGACATGAGGGCAGTGTGACACCCAGGCGAGGAAAG	3401
Db	2719	ACCACCAATGAGAGGCTGACCTGACATGAGGGCAGTGTGACACCCAGGCGAGGAAAG	2778
Qy	3402	AAGAAATGGCAGCCAGGCGACAGAGTCCGCACTCTCAGCTGTGGCCCTTGGTACAGCAC	3461
Db	2779	AAGAAATGGCAGCCAGGCGACAGAGTCCGCACTCTCAGCTGTGGCCCTTGGTACAGCAC	2838
Qy	3462	TGGGTGCCGGGGCA CGGCTGTGGAGAGCTGCGCAACAGCTGTGTGTGTGTGCTGCC	3521
Db	2839	TGGGTGCCGGGGCA CGGCTGTGTGGAGAGCTGCGCAACAGCTGTGTGTGTGTGCTGCC	2898
Qy	3522	TACACGGGTGCCATGACGACCTTCCGACACTCTTCCGAGAGCTAGACACCGGCTG	3581
Db	2899	TACACGGGTGCCATGACGACCTTCCGACACTCTTCCGAGAGCTAGACACCGGCTG	2958
Qy	3582	GCGGAGCTGAGGCTCACTGTGCTACGGGATCTCCGACACACAGCTCGAGAGATCTTCCTG	3641
Db	2959	GCGGAGCTGAGGCTCACTGTGCTACGGGATCTCCGACACACAGCTCGAGAGATCTTCCTG	3018
Qy	3642	AAGGTGTGAGGAGTGTGTGCGGACACAGATATGAGGATGGCAGCTGCGGGCAGCAC	3701
Db	3019	AAGGTGTGAGGAGTGTGTGCGGACACAGATATGAGGATGGCAGCTGCGGGCAGCAC	3078
Qy	3702	CTATGCAACGCAATTCCTGGCTAGAGCTAACCTCTGGCTCAAGATGCCGCCACAGGAG	3761
Db	3079	CTATGCAACGCAATTCCTGGCTAGAGCTAACCTCTGGCTCAAGATGCCGCCACAGGAG	3138
Qy	3762	ACAGCGCTGAGAGACCGGGGAACAGCTGGGTTCAGCCCCAGAGACTGACACAGGCTCTGGG	3821

Db	3139	ACAGCGCTGGAGAACGGGAAACAGCTGGGTTCAGCCCCAGAGACTGACACAGGCTCTGGG	3198
Qy	3822	CCAGACCGCGTGGGCGGGGTACAGGGCTGGGCACTGACCGCCAGCAGCTCCAGGCGCTG	3881
Db	3199	CCAGACCGCGTGGGCGGGGTACAGGGCTGGGCACTGACCGCCAGCAGCTCCAGGCGCTG	3258
Qy	3882	CTTCTCAAGCGCTTCTGTCTTGGCCCGCAGCGCGCGGGCCTGTGTGGCCAGATCGTG	3941
Db	3259	CTTCTCAAGCGCTTCTGTCTTGGCCCGCAGCGCGCGGGCCTGTGTGGCCAGATCGTG	3318
Qy	3942	CTGCTGCCCCCTTTGTGGGCCCTGTGGCTCGTGTTCAGGCTCATCTGTGCTCTCTTCGGG	4001
Db	3319	CTGCTGCCCCCTTTGTGGGCCCTGTGGCTCGTGTTCAGGCTCATCTGTGCTCTCTTCGGG	3378
Qy	4002	CACTACCGGGCTGTGGGCTCAGTCCACCATGTACGCTGCTCAGGTGCTCTTCTTCAGT	4061
Db	3379	CACTACCGGGCTGTGGGCTCAGTCCACCATGTACGCTGCTCAGGTGCTCTTCTTCAGT	3438
Qy	4062	GAGGACGCCCCCAGGGGACCTTGACGCTGCCCGGCTGCTCGAGGCGCTGCTGCAGGAGGCA	4121
Db	3439	GAGGACGCCCCCAGGGGACCTTGACGCTGCCCGGCTGCTCGAGGCGCTGCTGCAGGAGGCA	3498
Qy	4122	GGACTGAGGAGCCCCCAGTGTGACGATAGTCTCCACAGGTTCTCGGCACACAGAGTTCTCT	4181
Db	3499	GGACTGAGGAGCCCCCAGTGTGACGATAGTCTCCACAGGTTCTCGGCACACAGAGTTCTCT	3558
Qy	4182	GCTGAAGTGGCCAAAGTCTTGGCCAGTGCACACTGACCCAGAGTCTCCATCCCCAGCC	4241
Db	3559	GCTGAAGTGGCCAAAGTCTTGGCCAGTGCACACTGACCCAGAGTCTCCATCCCCAGCC	3618
Qy	4242	TGCCAGTGTAGCAGCAGCCCGCTGCCCGGCTGCTGCTCCGACACTGCCCCGCTGAGTGGT	4301
Db	3619	TGCCAGTGTAGCAGCAGCCCGCTGCCCGGCTGCTGCTCCGACACTGCCCCGCTGAGTGGT	3678
Qy	4302	GGTCCCCCTCCGCCCCCAGGACGAGTACCAGGCTCTGGGGAAGTGTTCAGAACCTTGACAGGC	4361
Db	3679	GGTCCCCCTCCGCCCCCAGGACGAGTACCAGGCTCTGGGGAAGTGTTCAGAACCTTGACAGGC	3738
Qy	4362	CGGAACTGTCTGACTTCTGCTTCAGACCTACCCCGGCTGCTGCGCCAGGCGCTGGAAG	4421
Db	3739	CGGAACTGTCTGACTTCTGCTTCAGACCTACCCCGGCTGCTGCGCCAGGCGCTGGAAG	3798
Qy	4422	ACTAAGAAATGGGTGAATAGGTACAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCA	4481
Db	3799	ACTAAGAAATGGGTGAATAGGTACAGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCA	3858
Qy	4482	GGCTGCCCTCCGCGCCCAAGAGTTGGCCGCTCAGTGGAGAGTTGTGGGCGCTGTGAGT	4541
Db	3859	GGCTGCCCTCCGCGCCCAAGAGTTGGCCGCTCAGTGGAGAGTTGTGGGCGCTGTGAGT	3918
Qy	4542	CCCCCTGGCGGGGGCCCTCGACCGTGTCTGAAATACTCACGCCCTGGGCTCACAGC	4601
Db	3919	CCCCCTGGCGGGGGCCCTCGACCGTGTCTGAAATACTCACGCCCTGGGCTCACAGC	3978
Qy	4602	CTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTG	4661
Db	3979	CTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTG	4038
Qy	4662	GCCTTTGTCAACCGAGCCAGCAACGAAATCTCCGCTGCTCACTGCCCCCAGGCGCGGCC	4721
Db	4039	GCCTTTGTCAACCGAGCCAGCAACGAAATCTCCGCTGCTCACTGCCCCCAGGCGCGGCC	4098
Qy	4722	CGGCACGCCACAGCATCACACACTCAACCCCTTGNACCTCACAGGAGGAGCTG	4781
Db	4099	CGGCACGCCACAGCATCACACACTCAACCCCTTGNACCTCACAGGAGGAGCTG	4158
Qy	4782	TCTGAGGCTGCACTGATGGCTCTCTCGGTGGAGCTCTCTGCTCTCCATCTGTGTGGTCTTT	4841
Db	4159	TCTGAGGCTGCACTGATGGCTCTCTCGGTGGAGCTCTCTGCTCTCCATCTGTGTGGTCTTT	4218
Qy	4842	GCCATGCTCTTTGTCCCGGCCAGCTTCACTTTGCTCTCATTTGAGGAGCGAGTCAACCGA	4901
Db	4219	GCCATGCTCTTTGTCCCGGCCAGCTTCACTTTGCTCTCATTTGAGGAGCGAGTCAACCGA	4278

Qy	870.	TCAGAGCCCTCTCAGTGTTCAGGGGACCTAGCAGCAGTGGGCCCTCCCTCAACTGG	929
Db	652	TCGGAGCCCTCTCAGTACCAAGGGGCCAGCAGTCCAGGGGTCTGTCCCTCAACTGG	711
Qy	930	TACAGAGCTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGANATCCGCCCTGCCAGAC	989
Db	712	TACGAAGCCAAACAGATCAACGAATTCATGGGGCTTGAGTGGCGCCACCCCTGCCTGAC	771
Qy	990	AGCAGCCTGAGCCCCCGCTCTCGGAGCTGATTTGGAGCCCTGGACACCCACCGCTGTCC	1049
Db	772	AGCAGTCTCAGCCCCCGCTCTCGAGTTCGTGGGGCAGCTGGATGACCAACCCCTGTGTCT	831
Qy	1050	CGCCTGCTCTGGAGACGCTTGAAGCCCTCTGATCTCGGGAAGCTACTCTTTTGACACAGAT	1109
Db	832	CGGCTGCTCTGGAGACGCTTGAAGCCATTCATCTCGGGAATTCCTTTTGACCTGAC	891
Qy	1110	ACACCTTTTACCCGGAAGCTCATGGCCAGGTGAACCGGACCTTTCGAGGAGCTCACCCCTG	1169
Db	892	ACAAATTCATAGGAACTCATGGCTCAGGTGAACCAAGACCTTTGAGGAGCTGGCTCTG	951
Qy	1170	CTGAGGAGTCTCGGGAGGTGTGGAGATGCTGGGACCCGGATCTTCACTTCATGAAC	1229
Db	952	CTGAGGACCTTACACGAACCTTGGGAGTACTGGGACCCAGATCTTCAACTTCATGAAT	1011
Qy	1230	GACAGTTCCAATGTGGCCATGCTCGACGGCTCTCGACAGATGCAAGATGAAGGAAGG	1289
Db	1012	GACAGTACAAAGTGGCCATGCTTCAGAACTGCTGGATGTGGAGGCAAGGGTGGCAG	1071
Qy	1290	CAGCCAGACCTGGAGCCGGGACCAATGAGAGCCCTGCGATCTTTTGGACCTCTGGG	1349
Db	1072	CAGCAGACACCCAAAGGCCAAGCAGTGTGAGGCTATCAGAGACTTTCTGGATCCTAGT	1131
Qy	1350	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCACCTGTGTGGGCAAGCTG	1409
Db	1132	AGGGTTCGATACAACTGGCAGGAGGCTCATGCAACATGGGGCGCTTGGCAGAAATACTG	1191
Qy	1410	GGCCGAGTGAAGGAGTGCCTGTCTTGGACAAGCTGAGGCGGCACCCCTCAGAGGCGAGCC	1469
Db	1192	GGCCAAATCTTGGAGTGTGTCTCTGGACAAGCTGAGGAGCTGTGCCCTCAGAGGAAGCT	1251
Qy	1470	CTGTGTGCGGGGCCCTTGCAATCTGTCTCGGGAACATCGATTTGTGGGCGCGCTGCTTTC	1529
Db	1252	CTTGTGTCCGTGCGCTGGAGCTGCTGGGTGAGCGCGCCCTCTGGGCAAGCATCTCTTC	1311
Qy	1530	TTGGGACCTGAGGACTTTTCAAGACCCACAGAGCACCCACCCCA---GACCTGGGCCCC	1586
Db	1312	CTGAGCCAGAGACCCCTCTGGACTCATCTGAACCGCCATCTCAACCAACGAGGCTCT	1371
Qy	1587	GGCCACGTGGCATCAAAATCCGATGGACATTTGACGTGTTCAGAGAGCAATTAAGATC	1646
Db	1372	GGCCACCTACAGTCAAGATTCGATTCGATTTGATGATGTCAAGAGCAATTAAGATC	1431
Qy	1647	AGGACAGGTTTTGGGACCCCTGGGCCAGCCGCGGACCCCTTGACCGACCTCGCTACGCTG	1706
Db	1432	AGGGACAAGTTTTTGGGACCCAGGTCCGTGACAGATCCCTCTCATGGACCTTCGTATGTG	1491
Qy	1707	TGGGGCGGCTTCGTGTACTTGCAAGACCTGTGTGAGGCGGTGAGCCGTCCGCGTCTCAGC	1766
Db	1492	TGGGGCGGCTTCGTGTACTTGCAAGACCTGTGTGAGCAGGACGTGTGCGCGTCTCTGAGT	1551
Qy	1767	GGGCCCAACCCCGGGCGGGCTCTACTGTGACAGATGCCCTATCCGTGCTATGTGGAC	1826
Db	1552	GGCAGGACTCCAGACAGGCGCTCTATTTTCAGACAGATGCCACCCCTGCTACGTGGAC	1611
Qy	1827	GACGTGTTCTCGTGTCTGAGCCGCTGCTGCGCTCTTCTCTGACGCTGGCCTGGATC	1886
Db	1612	GATGTGTTCTCGGGTGTGAGCCGCTCTGCGCTCTGTCTCTGTCTCTGACTCTGGCTGGATC	1671
Qy	1887	TACTCCGTGACACTGACAGTGAAGCCGCTGTGTGGGAGAGGAGACCGCGCTGCGGGAC	1946
Db	1672	TACTCGGTGGCGCTCACTGTGAAGGCGCTGTGTGGCGAGAGAGACAGCGCTGCGAGAA	1731

QY	1947	ACCATGCGCGC	CAATGGGGCT	CAGCCGCGCGGTGCTCTGGCTAGGCTGGTTCCTCAGCTGC	2006
DB	1732	ACCATGCTGCGAT	TGGGGT	TGAGCCGTCGTGCTCTGGCTTGGTTCCTCAGCTGC	1791
QY	2007	CTCGGGCCCTT	CCTGCTCAGGGCCGCGCTGCTGGTCTGGTGTCTCAGCTTGGGGGACATC	2066	
DB	1792	CTGGGACCCCTT	CTTTGGTCAGCGCTGCATTTGGTGTAGTGTCTTAAGCTAGGGAAATC	1851	
QY	2067	CTCCCTCTACAGCCAC	CCGCGCGTGTCTTCTCTGTCTTTGGCAGCCTTCGCGGTGGCCACG	2126	
DB	1852	CTTCTCTACAGCCAC	CCGGTGTAGTCTTCTCTTTCTTGGCGGCTTGGGTGGCCACC	1911	
QY	2127	GTGACCCAGAGCTT	CTGTCTCAGCGCTTCTTCTCCGCGGCCAACCTGGCTGGGGCTGC	2186	
DB	1912	GTGCTCAGAGTTTT	CTGTCTCAGTGCCTTCTTCTCCGCGGCCAACCTGGCAGCAGCTGC	1971	
QY	2187	GGCGGCTTGGCTACTT	CTCCCTCTACTTCGCCCTACGTGCTGTGTGTGGCTTGGCGGAC	2246	
DB	1972	GGGGGCTTAGCTACTT	CTGGGCTCTACTGCCCTATGTGCTGTGTGTGTGCTTGGCGTAG	2031	
QY	2247	CGGCTGCGCGGTGGCTG	CGCGCGCAGCCTGTCTGCGCCGTGGCGGCTTCGGCTTC	2306	
DB	2032	CGCTGCTCTGGCGGACTT	TAGCTGTGAGCTGCTGTCTCCCGTGGGCTTGGCTTT	2091	
QY	2307	GGCTGCGAGAGCTTGGCT	CTGTCTGGAGGAGCAGGGGCGGCGCAGTGGCACAACGTG	2366	
DB	2092	GGATGCGAGAGCTTGGCT	CTACTTGGAGGAGCAGGGGACGCGGCTCAGTGGCACAATTTG	2151	
QY	2367	GGACCCGGGCTTACGGCAG	AGTCTTCAGGCTGGCCCGCCAGTCTCTGGCTTCTGTGCTG	2426	
DB	2152	GGCACAGGCCCGGGAGS	ATGTCTTCAGGCTTGCCACAGGTGTCTGCTTCTCTTGTGCTT	2211	
QY	2427	GACGCGCGCTCTACGSCCT	CGCCACTGTGTACTGTGMACTGTGTGTGCCCAGAGCAGTAC	2486	
DB	2212	GATGCTGTCACTACGSCCT	TGCCCCTGTGTACTAGAGGCTGTATGCCCAGGCGCAGTAT	2271	
QY	2487	GGGATCCCTGAAACCAT	GGAAATTTTCTTTTTCGAGGAGCTACTGTGTGCGGACCTTGCGCCC	2546	
DB	2272	GGAACTCCCTGAAACCAT	GGAAATTTTCCCTTTTCGAGGAGCTACTGTGTGCGCCCTGGACCT	2331	
QY	2847	CCCAAGAGTTCAGCCCTT	GCACCACCCCGCTGGACCCCAAGGTGTGTGTAGAGAGGCA	2606	
DB	2332	CCCAAGAGTCTGTCTT	GGCCCTTGCCCCCAAGATCCCAAGGTCTGTGTGGGAGAGGCCA	2391	
QY	2607	CCGCGCGGCTGAGTCT	CTGGCGTATCCGTTTCGACGCTGGAGAGCGCTTTCCTGGGAAGC	2666	
DB	2392	CCACTGGTCTGGTTCT	CTGGTGTCTCAATTCGAGGCTGGAAGAGCAATTTTCTGGTCTCT	2451	
QY	2667	CCGAGCCAGCCTTGGCGG	GGCTCAGCCTTGGACTTTCACGAGGCCACATCAACGCTTC	2726	
DB	2452	CCACAGCCAGCCTTGGAG	GGCTCAACTTTGACTTCTATGAAGGCCACATCACTGCCTTT	2511	
QY	2727	CTGGGCCACAACGGGGC	GGGAAGACCAACACCTGTCTCAATCTTGTAGTGGCTCTTCCCCA	2786	
DB	2512	TTGGGTCAACAACGGGG	CTGGCAAGACAAACACACTGTCTCAATCTTGTAGTGGCTCTTCCCCA	2571	
QY	2787	CCAGTGTGTGCTCTGCT	CTTCACTCTGGGCCACAGAGCTCGCTCCAGCATGGCGCCATC	2846	
DB	2572	CCCAGTAGTGGCTCGG	CCCTCATCTGGGGCATGACGTGCAAAACCAACATGGCAGCCATC	2631	
QY	2847	CGGCCCCACCTGGGCGT	CTGTCTCAGTACAACGTGCTGTTTGAATGACATGTCGCGGAC	2906	
DB	2632	CGGCCCCACCTGGSCAT	CTGCCCAACAGTACAATGTGCTGTTTGTACATGCTGACGGTGA	2691	
QY	2907	GAGCAGCTCTGTTCTAT	TGGGGCGGCTGAAGGGTCTGAGTCCGCTGTGTGGGCCCCGAG	2966	
DB	2692	GAAACAGTTTGGTTCT	ATATGGCCGTTTGAAGGGGCTGAGTGCAGCCGCATATAGACTCCGAG	2751	
QY	2967	CAGACGCTCTGTGCAAG	GATGTGGGCTGTCTCCAAGCAGAGTGTGCAGACTCGCCAC	3026	
DB	2752	CAGGAAATTTGATTCGG	GAATGTGGGCTCATCCCTTAGCGGGACACACAGACGCTCAC	2811	
QY	3027	CTCTCTGTGGGATGCA	ACGGAAGCTGTCTCGGTGGCCATTCCTTTGTGGCGGCTCCCAA	3086	

GLPVGVSIRGLKXHPGCPQALQGLNLDVFEHGHITAPLGHNGAGKTTTILSILGLFP
PSSGSASILGHVDVQTNMAAIRPHLIGICQYNVLFDMLTVEHVMFYGRKGVSAAMG
PERELRILGKTRDQTRHRLGSGQMOKLSVAIAFVGGSRVIMDEPTAGVDPASR
RGIWELLKYEGRHTLILSTHLEDEALLDRVAMVAGSILCCSGSPFLRHLGCGV
YLTIVKSSQSILVTHDAKSDSDPREKSDGNGRTSDAFTRGSDSKSNOAPAGCAV
ITPSTARILELVQOHVPCAQLVEDLPHELLAVLPAAGALDGSFAMVQELDQOELLEL
LTVGISTNLEELFPAKVEDAHREGDSRRLQHLRITCTPPTGPEASVLENGELAK
LVLDOPAPQGLNMAAQVQWTLTCQOURLLHFRLLARRSRRLFAQVVLPAFLVFG
LALLFSLIVPFGYPIQLSPAMYPQVFSFSEADPGDPMRMLLEALLGEAGLOEP
SMODKARGSECTHSLACYFTVPEVPDVASILASGNWTPESPACQCSOPGARRLL
PDCPAGAGPPPPQAVAGLEVONLTGRNVSDFLVKTYPSLVRGLTKWKVDEVRY
GGFSIGRDPDLPFGHEVVRTLAEIRALLSPQGNALDRILNLTQWALGLDARNSLK
IFWNNKHAMVAFWNRANGLHALLPSPGVRHAHSITLLNHLNLTQOULSEATLI
ASSVDLVISICVWFAMSPASFTVLVLIBERITRAXHLQVLVSGLPQTLVQNLFLDM
CNYLVAIVVFIIFLAFOQYAVAPENLPALLLLLLYGNISITPLMTPASFFVSPST
AYVVLTCINLIFGINSMAFLVELLSQNLQEVSRILKQVFLIPPHFQGRGLIDMV
RNOAMDAFERLGDKOFOSPLRWDIIGNVLLAMMAQGLFLITLLOHNRLLPOSK
PRLLPPLGEEDVQAQREVRVTKGATQGDVLVRLDLYKVRGQRNPAVRLCLGIPPG
ECFGLLVNGAGKSTFRWVTGDTLPSSGEAVLAGHNVAQERSAAHRSMGYCPOSDAI
FDLLTGREHLELFARLGRVPEAQVAQTALSGVRLGLPSYADRPAQYSGNKRKLAT
ALALVGDPAVVFLDEPTTGMDSARRFLNLSLLSVREGRSVLTSMSHECEALCTR
LAIWNGRFRCLGSHQHLKGRFGAGHTLTLRVPDPQPEPAFAFIRIYFPGAELREVHG
SRLRFPPLPGRCITLTVRVELAAQRAHGVDEFSVQTTLEEVFLYFSKDOQGBEBS
SRQAEBEVSKPRQHPKRVSRFLEDPSSVETMI"
1. .138

sig_peptide
ORIGIN

Query Match 57.4%; Score 3902.4; DB 6; Length 6633;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 4991; Conservative 0; Mismatches 1456; Indels 135; Gaps 8;

Qy	210	ATGCGCTTCTGGACACAGCTGATGCTGCTGTGGAAGAATTTCAATGATCGCCGGAGA	269
Db	1	ATGCGCTTCTGCACACAGTTGATGCTTCTGCTGTGGAATAATTACACCTATCGACGGAGA	60
Qy	270	CAGCGGTCCAGCTCCTGGTGGAAATGCTGTGGCTCTCTTCTCTTCTTCTTCTTCTTCTG	329
Db	61	CAACCGATCCAACTACTAGTGGAGTGGCTTTTGGCCCTCTTCTCTTCTTCTTCTTCTAGT	120
Qy	330	GCTGTTGCCACTCCACCCGCTGGAGCACATGAATGCCATCTTCCCAACCAAGCCA	389
Db	121	GCTGTGCTGCTACTCCACCCGCTGGAGCATACGAAATGCCATCTTCCCAACCAAGCCA	180
Qy	390	CTGCCATCGCGGGACCGTCCCTGCTCGAGGTCTCATCTGTAATGTGAACAACACC	449
Db	181	TTACCATCGCGGGACCGTCCCTGCTGGAGTGGTCTGCTGCAACGTAAACACTCC	240
Qy	450	TGCTTTCGGAGCTGACACCGGGGAGAGCCCGGGCCCTGAGCACTTCAACGACTCC	509
Db	241	TGCTTTCAGCACCCAAACGCTCGGAGAAAGCTGGGGTCTGAGTAACCTTTAAGGATTC	300
Qy	510	CTGGTCTCCGGCTGCTAGCGGATCCCGCACTGTGCTGGAGGGGCGGCGAGTCCACAGG	569
Db	301	TTGATCTCGAGCTCCTCGCTGATACCCGCAAGTGTCTGGGGGGCCAGCATCCAGGAC	360
Qy	570	ACGCTGGCTGCCCTAGGGAAGCTGATCGCCACGCTGAGGGCTGACGAGCACGCGCCGAG	629
Db	361	ATGCTGATGCCCTGGGGAACTGATCCCGTGTCTCAGGCGAGTGGAGGTGGAGCACGA	420
Qy	630	CCTCAACCAACAGAGCTCTCCACTGGAACCAACCCATGCTGGATGTGCGGGAGCTGCTG	689
Db	421	CCACAGGAGAGTACACAGCCGACAGTCAAGGGTC-----AGTGACTTAAGCTTCTG	471
Qy	690	ACGTCACCTGCTGCGACAGGAATCCCTGGGGTTGGCACTGGGCCCAAGCCAGAGCCCTTG	749
Db	472	GAGAAGATCTTGCAAAGGGCATCTCTGATCCTGTGCTGGGTCAAGCCAGGATTCATG	531
Qy	750	CACAGCTTGTGGAGGCGCTGAGAACCTGGCCCGCAGGAGCTCTCTGGGCTCGCAGGCTG	809
Db	532	AGAAAGTTCTCAGATGCTATCAGGATCTTTCGCCAGGAGCTCTCTGACACTGCCAGGCTG	591
Qy	810	GTGAGCTTCCGGCACTGCTCAGAGACCCCGAGGACCCAGCGGCCCTCTGGAATGCTG	869

Db	592	ATGAGAGCTCCGAGCTTTTCTCTCGGAGGCCCCGAGAGGCTCAGCTGGTTCCTCTGGAGCTGGTT	651
Qy	870	TCAGAGGCGCTCTGCACTGTTCAGGGGACCTTAGCAGCACAGTGGGGCCCTCCCTCAACTGG	929
Db	652	TCGAGGCGCTCTGCACTGTACCAAGGACCCAGCAGTCCAGGGGGCTGTCCCTCAATTGG	711
Qy	930	TACAGGCTAGTACCTGATGAGAGCTGTGTGGGAGGAGCCAGAAATCCGCTCTGCAGAC	989
Db	712	TACGAAGCAACACAGCTTAATGAGTTATGCTGAGGCTGAGAGTGGCCCTGCTGCTGCTGAC	771
Qy	990	AGCAGCTGAGCCCGCTGCTCGGAGCTGATGGAGCCCTGGACAGCCACCCGCTGTCC	1049
Db	772	AACAGCTCAGCCCTGCTGCTGCTGAGTTTGTGGGACCTGGATGACCACTGTGTCT	831
Qy	1050	CGCTGCTCTGGAGACGCTGAAGCCTCTGATCTCTCGGAAAGCTACTCTTTCGACACAGAT	1109
Db	832	CGCTGCTCTGGAGGCGCTGAAGCCTGATCTCTCGGAAATTCCTTTCGACCTGAC	891
Qy	1110	ACACCTTTTACCCGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAGCTCACCTG	1169
Db	892	ACAACTTCACTCGGAAGCTCATGGCTCAGGTGAACAGACCTTCGAGGAGCTGGCTCTG	951
Qy	1170	CTGAGGATGCTCGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACCTTCATGAC	1229
Db	952	TTGAGGGACCTTACACGAACCTCTGGGGGTGCTGGGACCCGAGATCTTCAACTTCATGAAT	1011
Qy	1230	GACAGTTTCCAATGTGGCCTGCTCAGCGGCTCTCGAGATGAGGAGTGAAGGAAGG	1289
Db	1012	GACAGTACCAAGTGGCCATGCTTCAGAGGCTTCGTGATGTGGGGGACAGGGCAGAG	1071
Qy	1290	CAGCCACAGCTGGAGCCGGGACCAATGAGGCGCTCGCATCTTTCGAGCCCTGG	1349
Db	1072	CAGCAGACACCCAGAGCCAGAAAGTTGGAGCTTCAAAGACTTCTTGATCCTAGT	1131
Qy	1350	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGACCTGGTGGGACGCTG	1409
Db	1132	AGGGTGGCTACAGCTGGCGGAGGCCACGAGATATGGGACGCTGGCTGGAAATCTTA	1191
Qy	1410	GGCGAGTGAAGAGTGCCTGCTGGACAAGCTGAGGGGCGACCTCAGAGGACAGCC	1469
Db	1192	GGACAAATGATGAGTGTGTCTCGACAAAGCTGAGGCTGTGCCCTCAGAGGAGCT	1251
Qy	1470	CTGTGTGGGGCCCTGCAATGCTCTCGGAAACATCGATTCGTGGGCGGCGTCTGCTTC	1529
Db	1252	CTTGTGTCCGTGCGCTGGAGCTGCTGGGTGAGCGCCCTCTGGCGAGGACATCGTGTTC	1311
Qy	1530	TTGGGACCTGAGGACTCTTCAGACCCACAGAGCACCCAAACCCAGACCTGGGCGCCG	1589
Db	1312	CTGAGCCAGAGCATCTCTGGACCCCATCCGAACCTGTCTCTCCAGCCCTGAGTCTGGC	1371
Qy	1590	CACGTGGCATCAAAATCCGCATGGACATTTGACGTGTGTCAGAGGACCAATTAAGATCAGG	1649
Db	1372	CACCTAGATTCAGATTCGAATGGATATCGATGATGTCAAGAGACCAATTAAGATCAGG	1431
Qy	1650	GACAGTTTTTGGACCCCTGGCCAGCGCCGACCCCTGACCCGACCTGCGCTACGTGTGG	1709
Db	1432	GACAAGTTTTGGGACCCAGGTCGCTCAGCAGATCTCTTTCATGGACCTTCGCTATGTGG	1491
Qy	1710	GGCGCTTCTGTACTGCAAGACCTGTGAGGCTGCGAGCCGCTCGGCTGCTCAGCGGC	1769
Db	1492	GGAGCTTCTGTGTAACCTGCGAGGACCTGCTGAGCAGGCTGTGCGAGTGTCTGGTGGC	1551
Qy	1770	GCCAAACCCCGGGCGGCTCTTACTCTCAGCAGATGCCCTATCTCGTGTATGTGGACGAC	1829
Db	1552	GGNACTCCCGCAAGGTCTTACTCTGAGCAGATGCCACCCCTGCTGCTGCTGATGAT	1611
Qy	1830	GTGTTCTGCTGTGTGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1889
Db	1612	GTGTTCTGCGGGTGTGAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1671
Qy	1890	TCGTGACACTGACAGTGAAGCCGCTGTGGGAGAGGAGCGGCTGCGGAGACACC	1949
Db	1672	TCGTGCGCTCACTGTGAAGCCGCTGTGCTGTGAGAAAGAGACAGCGCTCGGAGAAC	1731

QY	1950	ATGGCGCATGGGCTCAGCGCGGCTGCTCTGGCTAGGCTGGTTCCTCAGCTGCCTC	2009
Db	1732	ATGCGTGCATGGGCTGAGCGCGGCTGCTCTGGCTTGGTGGTTCCTCAGCTGCCTG	1791
QY	2010	GGGCGCTTCTCTGCTCAGCGCGGCTGCTCTGGCTCAAGCTGGGCGGACATCCCTC	2069
Db	1792	GGACCTTCTCTGCTCAGCGCTGGTGGTCTGATATTAGCTTAGGGAACATCCCTT	1851
QY	2070	CCCTACAGCCACCGCGGCTGGTCTTCTCTGTTCTTGGCAGCTTCGGCGTGGCCAGGGTG	2129
Db	1852	CTTACAGCCACCGGCTGTAATCTTCTTCTTCTGGCGGCTTGGCGTGGCCACCGTC	1911
QY	2130	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAACCTGGCTGGGCTGGGC	2189
Db	1912	GCTCAGAGTTTCTGCTCAGCGCTTCTTCTCCAGGGCCAACTGGCGAGCAGCTCGGG	1971
QY	2190	GGCTGGGCTTCTTCTCCCTTACCTGCTTACGCTGCTGCTGCTGGCTTGGCGGGACCGG	2249
Db	1972	GGGCTGGCTTATTTGGGCTCTATCTGCCCTAGTACTGCTGCTGGCTGGCGGAGCGC	2031
QY	2250	CTGCCCGGGTGGCGGCTGGCGGAGCTGCTGCTGCCCGCTGGCTTGGCTTGGC	2309
Db	2032	CTGCACCTGGCGGACTCTTAGCTGGAGCTGCTGCTCCCTGTAGCCTTGGCTTGGGA	2091
QY	2310	TGGAGAGCTGGCTCTGCTGGAGGAGCAGGCGGAGCGGCGGAGTGGCAACGTTGGC	2369
Db	2092	TGCGAAGCTGGCGCTTACTAGAGGAGCAGGAGAGCGGGCTCAGTGGCACAATTTGGC	2151
QY	2370	ACCGGCTACGGCAGACGCTTTCAGCTGGCGCCAGGCTCTGCGGCTTCTGCTGGAC	2429
Db	2152	ACAGGCCCCGGAGGAGCTTTCAGGCTGGCGCCAGGCTGCTGCTTCTGCTTGTAT	2211
QY	2430	CGCGGCTTACGGCTTCGCCACTTGTAACCTGGAAGCTGTGCTCCAGGCCAGTACGGG	2489
Db	2212	GCGTCTATCTACGGCTTGGCTCTGGTACTAGAGGCTGTGTGCGCCAGGCAAGTATGGA	2271
QY	2490	ATCCTGAACCATGGAAATTTCTTTTCGAGGAGCTACTGCTGGGAGCTCGGCGCC	2549
Db	2272	ATCCCTGAACCATGGAAATTTCCCTTTTCGAGGAGCTACTGGTGTGGACCTCGGCTCC	2331
QY	2550	AAGAGTCCAGCCCTTGGCCACCCGCTGGACCCAAAGGCTGTGTAGAGAGGACCG	2609
Db	2332	AAGATTCTCTTGGCCCCCTGCCCAAGATCCCAAGGTTCTGGTGAAGAGCACCC	2391
QY	2610	CCCGGCTGAGTCTGCTGGCTATCCGTTTCGAGCTGGAGAGCGCTTCTTGAAGCCCG	2669
Db	2392	CTTGGGCTGGTCTGCTGGTCTCCATTCGAGGCTGGAAGAACATTTCTGGCTGTCCG	2451
QY	2670	CAGCAGCCCTCGGGGGCTCAGGCTGGACTTCTACAGGGGCCACATCAGCGCTTCTTG	2729
Db	2452	CAGCCAGCCCTGCAAGGACTCAACCTTGACTTCTACGAAGGCCACATCACTGCTTTTG	2511
QY	2730	GGCCAAACCGGGCGGCAAGACCACCACTGTGCTCATCTTGAGTGGCTCTTCCACCC	2789
Db	2512	GGTCAACCGGGCTGGCAAGAACCACTGTCTCATTTTGAGTGGTCTTCTTCCACCC	2571
QY	2790	AGTGGTGGCTCTGCTTTCATCTTGGGCCACGAGCTCGCTCCAGCATGGCGCCATCCGG	2849
Db	2572	AGTAGGGCTCGGCTTCCATCTTGGGCTATGATGTACAAACCAACATGGCAGCATCCGG	2631
QY	2850	CCCGACTGGGCTGCTGCTCTCAGTACAACTGCTGTTTGAATGTGACATGTGACCTGGAG	2909
Db	2632	CCCGACTGGGCTGCTGCTGCGGAGTACAATGTGCTGTTGATATGTGACAGTGGAGAA	2691
QY	2910	CAGCTGTGTTCTATGGCGGCTGAAGGCTGAGTGGCTGAGTGGCGCCCGGAGCAG	2969
Db	2692	CATGTTTGGTCTATGAGCGTTTGAAGGGGTGAGTGCAGCGGCTATGGCCCCCGAGCGG	2751
QY	2970	GACCGTCTGCTGAGGATGTGGGCTGGTCTCCAAAGCAGAGTGTGACAGTTCGCCACCTC	3029
Db	2752	GAACGTTGATACGGGATGTGGGCTTACCTCAAGGGGAGCACACAGACACGCCACCTC	2811

QY	3030	TCTGTTGGATGCAACGGAAGCTGTCCTGTGGCCATTGCTTGTGGCGGCTCCCAAGTT	3089
Db	2812	TCTGTTGGAATGCAAGAAACCTTCTGTGGCCATTGCTTGTGGTGGCTCTCTGTGTG	2871
QY	3090	GTTATCTCTGAGCAGCTTACGGCTGGCTGGATCTCTGTCTCCCGCGCGGTATTTGGAG	3149
Db	2872	GTCACTATGAGCAGGCCACTGCTGTGTGGACCCGCTTCCCGCGTGGCAATTTGGAA	2931
QY	3150	CTGCTCTCAATATCCGAGAAGTCCGACGCTGATCTCTCCACCCACCACTGGATGAG	3209
Db	2932	TTGCTACTTAAAGTACAGAGAAGTCCGACACTGATTTCTCTCCACTCACCTGGATGAG	2991
QY	3210	GCAGAGCTGCTGGAGACCGTGTGGCTGTGGCAGGTGGCGCTTGTCTCTGTGGC	3269
Db	2992	GCAGAGCTCTTGGAGATCCGTTGGCCATGTGGCAGTGGCTTGTGTGTGTGGTGG	3051
QY	3270	TCCCCACTCTTCTGCGCGCTCACTTGGGCTCCGGCTACTTACCTGACGCTGTGTAAGGCC	3329
Db	3052	TCCCCGCTTTTCTTGGCGGACACTTGGGCTTGGGTTACTACCTGACCTCGTGAAGAGT	3111
QY	3330	CGCTGCCCTGACCAACAATGA-----	3352
Db	3112	TCTCAGTCTCTCTGTCACCCATGATGCTAAAGGGAGACAGTAGGACCCACAGACGGGAAAG	3171
QY	3353	-----GAAGGCTGACACTGACATGAGGGGAGCAGTAGGACCCACAGCAGG	3396
Db	3172	AGTCAAGTGGCAATGCGAGGAGCTCAGACACAGCGTTTACACAGAGAACCTCAGACAAAG	3231
QY	3397	AAAA-----GAAGAAATGGCAGCAGGCGCAGAGTCCGCACTCTCTCAGCTGCTG	3446
Db	3232	AGCAACAGGCGCGGCTCTCTGGCGGCTTCCCATCACCCAGCACAGCCCGGATACTA	3291
QY	3447	GCCCTGGTACAGCACTGGGTGCCCCGGGACAGGCTGGTGGAGAGCTGCCACAGAGCTG	3506
Db	3292	GAGCTAGTGCAGCAGCATGTGCTGGAGCACAACTCGTGGAGACCTTGCCTCAGTACT	3351
QY	3507	GTGCTGCTGCTGCCCTACACAGGCTGCCATGAGCGGAGCTTGCACACACTTCTCCAGAG	3566
Db	3352	CTGCTTGTGCTACCTTATGCGGGGCTTGGATGGAGCTTCCGCAATGCTTTCAGGAG	3411
QY	3567	CTAGACACCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCCTC	3626
Db	3412	CTGGATCAGCAGCTGGAGCTCTTGGGCTCTCACAGGCTACGGGATCTCGGACACCAACTG	3471
QY	3627	GAGGAGATCTTCTTGAAGTGTGGAGGAGTGTGCTGGGACACACATATGGAGATGGC	3686
Db	3472	GAGGAGATCTTCTTAAAGGTGTGGAGGA-----TGGCAGACAGAAAGTGGGACTCT	3525
QY	3687	AGCTGGGCGAGCACCTATGCAAGGCAATTGCTGGCTTAGACGTAACCTTGGGCTCAAG	3746
Db	3526	AGACCGAGCTGCACCTTGCACATGC-----ACTCCACAGCCCCC	3567
QY	3747	ATGCCGCCACAGAGACAGCGCTGGAGAACGGGGAACAGCTGGGTGAGCCCCAGAGAC-	3805
Db	3568	ACAGGSCCAGAGGCATCAGTTCTGGAGATGGGAGCTGGCTAAGCTGGTGGATGCC	3627
QY	3806	--TGACACAGGCTCTGGGCGAGCGCGTGGGCGGCTGACAGGCTGGGCACTGACCCGC	3863
Db	3628	CAAGCCCCACAGGCTTGGCACCAACGCTGCCCAAGTGCAGGCTGGACCTTACCTGT	3687
QY	3864	CAGCAGCTCCAGGCGCTTCTTCAAGCGCTTCTGCTTGGCGCGCAGCGCCCGCGGC	3923
Db	3688	CAACAGCTCGGCTCTGCTTCCAAAGGTTTCTGCTTGTCTGCGCGAGCGCGCGGCG	3747
QY	3924	CTGTTTCCCCAGATCGTGTGCTGCTTCTTGTGGGCTGGCCCTCTGTTTTCAGGCTC	3983
Db	3748	CTGTTTGCACAGGTTGTTGTGCTGCTTGTGGGCTGGCCCTGTCTTTCAGGCTC	3807
QY	3984	ATCGTGCCTCTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCCACTGATGACGCTGCT	4043
Db	3808	ATTGTGCTCTCTTTCGGCAGTACCCACCCCTGACGCTCAGCCCTCTATGATGCGCT	3867
QY	4044	CAGGTGCTCTTCTTTCAGTGGAGGACGCCCGAGGAGCCCTGGAGCTGCGGCTGCTCGAG	4103


```
Qy 6231 AGCCGCAACATCTCAAGGCGAGATTCCGCGGCTCACACACTGACCTCGGGTGCCC 6290
Db 6088 AGCTCTCAGCATCTCAAGAGGAGGTTTCGGGGCTGGCCACACACTGACTCTAAGGGTCCCA 6147
Qy 6291 GCGCAAGGTTCCAGCGCGGAGCGGCTTCGTGCGGCCGAGTTCCCTGCGGTGCGAGCTG 6350
Db 6148 CCGGACAGCTGAGCGCGGATAGCTTTCATCAGGATCATTCTCCCTGGGGCTGAACTC 6207
Qy 6351 CGGAGGCAATGAGGCGCGCTGCGTTTCAGCTGCGCGCGGAGGCGCTGCGCGCTG 6410
Db 6208 CGGAGGTCACGCGACCGCTGCGCTTCCAACTGCCACCGGGGCGAGATGACCCCTG 6267
Qy 6411 GCGGCGCTTTGGAGAGCTGGCGGTGCCAGCGCGCAGACACCGCTGGAGGACTTTTC 6470
Db 6268 ACACGAGTGTTCAGGGAGCTGGCTGCCAGCGGAGGCGCCACCGTGGAGGACTTCTCT 6327
Qy 6471 GTGAGCCAGACGATGCTGGAGGAGTATCTTGTACTTCTCCAGGACACAGGGGAGGAC 6530
Db 6328 GTGAGCCAGACCACTCTGGAGGAGTGTTCATATATTTCTCCAAAGACCAAGGGAGAG 6387
Qy 6531 GAGGACACCAAGAGCAGAGAGGAGGAGGAGTGGAGTGAGACCCCGCGCAGGCTGCGAG 6590
Db 6388 GAAGAGAGC--AGTCGCGAGGAGCTGAAGAGAGGAGGTTTCCAAACCTGGCCGCGAG 6444
Qy 6591 CACCCGAAACGCTGACCGAGTTCCTCGATGACCTAGACTGCGGAGACTGTGCTCTGA 6650
Db 6445 CATCCCAACGCTGACGCGGATCTCTGGAAGACCCAGCTCTCTGGGAGACCATGATCTGA 6504
Qy 6651 GCCTCCCTCCCTGCGGGCGCGGGAGGCGCTTGGGAATGG 6692
Db 6505 GCATGCTGCTTGGGACTGAGTGGCAAGCTAGACAGAGG 6546

RESULT 15
AX080493
LOCUS AX080493
DEFINITION Sequence 32 from Patent WO0109314.
ACCESSION AX080493
VERSION AX080493.1 GI:13162147
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Chimini,G.
TITLE Novel uses of abca-type transporters
JOURNAL Patent: WO 0109314-A 32 08-FEB-2001;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE; (INSERM)
(PR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
source
1. .6607
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 57.1%; Score 3885.8; DB 6; Length 6607;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 4984; Conservative 0; Mismatches 1442; Indels 153; Gaps 8;

Qy 210 ATGGCTTTGGACACAGCTGATGCTGCTCTGGAAGAATTTCAATGATTCGCGGAGA 269
Db 1 ATGGCTTTAGGACACACAGCTGATGCTCTGCTGGAAGAAATTTACACCTATCGACGGAGA 60
Qy 270 CAGCGGTCCAGCTCCCTGGTTCGATTCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 329
Db 61 CAACCGATCAACACTACTAGTGGAGTGTCTTTGGGCGCTCTCTCTCTCTCTCTCTCTCT 120
Qy 330 GCTGTTCCGCACTCCCAACCGCGCTCGGAGCACCATGAATGCCACTTCCCAACAGGCCA 389
Db 121 GCTGTCGTCACCTCCCAACCGCGCTCGGAGCATCAGAAATGCCACTTTCACACAGCCA 180
```

```
Qy 390 CTGCCATCGCGGCGACCGTGCCTCGCTCCAGGGTCTCATCTGTAATGTGAACAACACC 449
Db 181 TTACCATCGCGGCGACCGTGCCTCGCTCGAGGGCTTGTCTGCAACGTAAACAACATCC 240
Qy 450 TGCTTTCCGAGCTGACACCGGCGCAGGAGCCCGGCGCTCGAGCAACTTTCAACGACTCC 509
Db 241 TGCTTTCCAGCACCAACCGCTGCGGAGAGCGCTGGGGTCTCTGAGTAATCTTTAAGGATTC 300
Qy 510 CTGGTCTCCGGCTGTAGCCGATGCCGACTGTGCTGGGAGGGGCGAGTGCCTCCACAG 569
Db 301 TTGATCTCGAGGCTCTCTGCTGATACCCGACAGTGTCTGGGGGCGCACAGCATCCAGGAC 360
Qy 570 ACCTGGCTGGCTAGGAGCTGATCGCCACGCTGAGGGCTGCAACGACACGCGCCAG 629
Db 361 ATGCTGGATGCGCTGGGAAACTGATCCCGCTGCTCAGGGCAGTTGGAGGTGAGACGCA 420
Qy 630 CTTCAACCAACCAAGCAGTCTCCACTGGAAACCAACCATGCTGATGCTCGCGAGCTGCTG 689
Db 421 CCACAGAGAGTACACGACCGGACCACTCAAGGTC-----AGTGACTAAGCTTCTG 471
Qy 690 ACCTCACTGCTGCGCACGGAATCCCTGGGGTTGGCACTGGGCCAAAGCCAGAGGCCCTTG 749
Db 472 GAGAAGATCCTGCAAGGGCATCCCTGGATCCTGTGCTGGTCAAGGCCAGAGATTCTATG 531
Qy 750 CACAGCTTGTGGAGGCGCTGAGGACCTGGGCCCAGGAGCTCTCTGGGCTGCGCAGCCTG 809
Db 532 AGAAAGTTCTCAGATCTATCAGGGATCTTGCCCGAGGAGCTCTGACACTGCCCAGCCTG 591
Qy 810 GTGAGCTTTCGGGCACTGCTGACAGACCCCGAGGAGCCAGCGGCGCCCTCGAGTGTGCTG 869
Db 592 ATGAGCTCCGAGCTTGTGCTGGAGGCCCGCGAGGCTCAGCTGGTCTCTGAGAGCTGGTT 651
Qy 870 TCAGAGGCTCTGCAAGTGTCAAGGCACTTAGCAGCACAGTGGGCGCCCTCCCTCAACTGG 929
Db 652 TCGAGGCGCTCTGCAAGTGTCAAGGCACTTAGCAGCACAGTGGGCGCCCTCCCTCAACTGG 711
Qy 930 TAGGAGCTAGTCACTGATGAGCTGCTGGGCGAGGAGCCAGAAATCCGCGCTCGCCAGAC 989
Db 712 TACGAAAGCAACCAAGCTTAATGAGTCTATGGGGCCAGAGTGGCGCTCGCCCTCGCTGAC 771
Qy 990 AGCAGCTGAGCGCCCGCTGCTCGAGCTGATTTGGAGCCCTGACAGCAGCCCGCTGTCC 1049
Db 772 AACAGTCTCAGCCCTGCTGCTGAGTTGTGGGACACTGGATGACCACTCTGTGTCT 831
Qy 1050 CGCTCTCTGGAGAGCGCTGAAAGCTCTGATCCTCGGAAGACTACTTTTGCACAGAT 1109
Db 832 CGCTCTCTGGAGGCGCTGAAAGCTTGTGATCCTCGGAAAAATTTCTTTGCACCTGAC 891
Qy 1110 ACACCTTTTACCGGAGCTCATGGCCCGAGTGAACCGGACCTTCGAGGAGCTCACCTG 1169
Db 892 ACAAACTTCACTCGGAAGCTCATGGCTCAGGTGAACCAAGACCTTCGAGGAGCTGGCTG 951
Qy 1170 CTGAGGAGTGTCCGGAGGCTGTGGGAGATGCTGGGACCCCGGATCTTCACTTTCATGAAC 1229
Db 952 TTGAGGAGCCTACACGAACTCTGGGGGTGCTGGGACCCCGAGATCTTCACTTTCATGAAT 1011
Qy 1230 GACAGTTTCAATGTGCCATGTGTCAGCGGCTCCTCGAGATGAGGATGAAGAAAGG 1289
Db 1012 GACAGTACCAACGTGGCCATGCTTTCAGAGGCTTCTGGATGTGGGGGCGCACAGGCGAGG 1071
Qy 1290 CAGCCAGACCTGGAGCGCGGACCACTGAGGCGCTTGGATCTCTTCTGGACCTGGG 1349
Db 1072 CAGCAGACCCAGAGCCAGAAAGTGGAGGCTTATCAAGACTTTTCTGGATCTCTAGT 1131
Qy 1350 AGCGGTGGCTTACAGCTGGCAGGACGACACGCTGATGTGGGCGACCTGGTGGGCAAGCTG 1409
Db 1132 AGGGTGGCTACAGCTGGCGGAGGCCACGACAGATATGGAGCGCTTGGCTGGAATCCTA 1191
Qy 1410 GGCCGAGTGAAGGAGTGTCTTGTGCAAGCTGGAGGCGGCACTCTCAGAGGAGCC 1469
Db 1192 GGACAAATGATGAGTGTGTCTCTCGGACCAAGCTGGAGGCTGTGCCCTCAGAGGAAGCT 1251
```


QY 1470 CTGGTGTGCGGGCCCTGCAACTGCTCGGGAACATCGATTCTGGGCGGCGCTGCTCTTC 1529
Db 1252 CTTGTGTCCCGTGGCTGGAGTGTGGGTGCGCGCGCTCTGGGCGAGCATCGTGTTC 1311
QY 1530 TTGGGACCTGAGGACTCTTTCAGACCCACAGAGCACCCAAACCCAGACCTGGGCCCCGGC 1589
Db 1312 CTGAGCCACAGACATCCTCTGAGCCCATCCGAACTGTCTATCTCCAGCCCTGAGTCTCTGGC 1371
QY 1590 CACGTGCGCATCAAAATCCGATGGAATTTGACGTGGTCAAGGACCAATAAGATCAGG 1649
Db 1372 CACCTACGATTCAAGATTGGAATGATATCGATGATGTACAAAGGACCAATAAGATCAGG 1431
QY 1650 GACAGGTTTGGGACCTGGCCAGCGCGGACCCCTGACCGACCTGGCGCTACGTGTGG 1709
Db 1432 GACAAAGTTTGGGACCCAGGTCCTGAGCAGATCCCTTTTCATGGACCTTCGGGTATGTGG 1491
QY 1710 GCGGCTCTGCTACTCTGCAAGACCTGGTGGAGCGTGCAGCGCTCCGGTGTCTCAGCGGC 1769
Db 1492 GGAGGCTTGTGTACTCTGAGGACCTGTGGAGGAGGAGCTGTGCGAGTGTCTCGGTGGC 1551
QY 1770 GCCAACCCCGGCGGCTCTTACCTGACGACGATGCCCTATGCCGTGCTATGTGGACGAC 1829
Db 1552 GGGAACTCCCGCACAGGTCTCTACCTGACGACGATGCCACACCCCTGCTACGTGATGAT 1611
QY 1830 GTGTTCTGCTGTGCTGAGCGGTGCTGCGGCTCTTCTGAGCCTGGCTGTGATCTAC 1889
Db 1612 GTGTTCTGCTGCGGTGCTGAGCGGTGCTGCGGCTCTTCTGACTCTGGCTGTGATTTAT 1671
QY 1890 TCCGTGACACTGACAGTGAAGSCCTGTGGGAGAGGAGACGCGGTGCGGACACC 1949
Db 1672 TCGTGGGCTCACTGTGAGGCCGTGTGGTGGAGAGGAGACGCGGTGCGAAGACC 1731
QY 1950 ATGCGCGCATGCGGCTCAGCGCGGCTGCTCTGGTAGGTGCTGCTCAGCTGGCTC 2009
Db 1732 ATGCGTGCATGCGGCTGAGCGCGGCTGCTCTGGTTGGTGTCTCAGCTGGCTG 1791
QY 2010 GGGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2069
Db 1792 GGAACCTCTCTGCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
QY 2070 CCTACAGCACCCGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2129
Db 1852 CCTACAGCACCCGGTGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1911
QY 2130 ACCAGAGCTCTCTCTCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2189
Db 1912 GCTCAGAGTTCTCTCTCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1971
QY 2190 GGCCTGGCTACTTCT 2249
Db 1972 GGGCTCGCTATTTTGGCGCTCTATCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2031
QY 2250 CTGCGCGGGTGGCGGCTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2309
Db 2032 CTGCACTTGGGCGGACTCTTACTGCTGAGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 2091
QY 2310 TCGAGAGCTCTGCTCTGAGGAGCAGGCGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCT 2369
Db 2092 TCGGAAAGCTTGGCGCTACTAGAGAGCAGGAGAGCGGCTCTGCTGCTGCTGCTGCTG 2151
QY 2370 ACCGGGCTTACGGCAGACGTCTTACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2429
Db 2152 ACAGGCCCGGGAGGAGCTCTTACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2211
QY 2430 GGGGCTCTACGGGCTCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
Db 2212 GCGGCTACTACGGGCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2271
QY 2490 ATCCCTGAACATGAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2549
Db 2272 ATCCCTGAACATGAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2331
QY 2550 AAGAGTCCAGCCCTTGGCCCAACCCCGCTGGACCCAAAGGTGCTGCTGAGAGGSCACCG 2609

Db 2332 AAGAGTCTGTCTGGGCCCTGCCCAACAAGATCCAAAGGTTCTGGTGAAGAGCCACCC 2391
QY 2610 CCGGCTGAGTTCCTGGGCTATCGGTTGCGAGCCTGGAGAGCGGCTTTCTCTGAGAGCCCG 2669
Db 2392 CTTGGCTTGGTCTCTGGGTCTCCATTCGAGGCTGAAGAACATTTTCTGCTGCTGCTG 2451
QY 2670 CAGCAGCCCTGCGGGGCTCAGCCTGGAATTTCTACGAGGCGACATCACCGCCTTCTCTG 2729
Db 2452 CAGCAGCCCTGCAAGGACTCAACCTTGACTTCTACGAAGGCCACATCACCTGCTTTTGG 2511
QY 2730 GGCACACGGGGCGGCGAAGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2789
Db 2512 GGTCAACAAGGGGCTGGCAAGCAACCACTGCTCATTTTGAAGTGTCTCTTCTCCACCC 2571
QY 2790 AGTGTGCTCTGCTGCTCATCTGCGGCCACAGAGCTGCTCAGCATGCGCGCCATCGG 2849
Db 2572 AGTAGTGTCTGCGCTCTCATCTGCGGCCATGATGTAACAACAATGCGAGCCATCGG 2631
QY 2850 CCCCACCTGGGCGTCTGCTCTCAGTACAAACGTGCTGTTTGAATGCTGACCTGAGCGAG 2909
Db 2632 CCCCACCTGGGCACTGCGCGCAGTACAAATGCTGCTGTTGATATGCTGACAGTGAAGAA 2691
QY 2910 CAGCTCTGCTTCTATGGCGGCTGAAGGCTCTGAGTGCCTGCTGAGTGGGCGGCCAGAG 2969
Db 2692 CATGTTGGTCTTATGGCCGTTTGAAGGCGTGAAGTGCAGCGCCATGGGCGCCGAGCG 2751
QY 2970 GACGCTGCTGCGAGGATGCGGGCTGCTCCTCAAGCAGAGTGTGCAGACTGCGCACCTC 3029
Db 2752 GAACTCTGATACGGGATGTGGGCTTACCTCAAGCGGAGACACACAGACGCGCACCTC 2811
QY 3030 TCTGTGGGATGCAACGGAAGCTGTCTCGTGCCCATTTGCTTTGGGGGGCTCCCAAGTT 3089
Db 2812 TCTGTGGAAATGACAGAGAAACTTCTGTGGCCATTTGCTTTGTGGTGGCTCTCTGCTG 2871
QY 3090 GTTATCTGAGAGAGCTACGGCTGCGCTGAGTCTCTGCTTCCGCGCGGCTATTTGGAG 3149
Db 2872 GTCATATGAGAGAGCCACTGCTGTGTGAGACCCGCTTCCGCGCTGGCATTTGGGAA 2931
QY 3150 CTGCTGCTCAAAATACCGAAGGTGCGACGCTGATCTCTCCACCCACCTCGATGAG 3209
Db 2932 TTGCTACTTAAGTACAGAGAAGGTGCGACATGATCTCTCCACTCAACCTCGATGAG 2991
QY 3210 GCAGAGTCTGCGGAGACCGTGTGCTGTGTGCGAGTGGCGGCTTGTGCTGCTGCTG 3269
Db 2992 GCAGAGCTCTTGGGAGATCGCGTGCCTCATGTGTGAGAGTGGCTCTTTGTGCTGCTG 3051
QY 3270 TCCCACTCTTCTGCGCGCTCACTGCGGCTTCCGCTCTACTCTGAGCCTGCTGAGGCG 3329
Db 3052 TCCCGCTTTTCTTGGCGCGACACTTGGGCTGCGGTTACTACTGACCTGCTGAGAGT 3111
QY 3330 CGCTGCGCTGACCAATGA----- 3352
Db 3112 TCTCAGTCCCTCGTCAACCATGCTNAGGGAGACAGTAGGAGCCCGAGAGGAGAAAG 3171
QY 3353 -----GAAAGCTGACATGATGAGGAGGAGTGTGAGACACACGAGG 3396
Db 3172 AAGTCAGATGCAATGGCAGGACGCTCAGACACAGGTTTCAACGAGGAACTCTCAGACA 3231
QY 3397 AAAA-----GAGAAATGGCAGCGGAGCAGAGTGTGGCACTCTCTCAGCTGCTG 3446
Db 3232 AGCAACAGGCGCGGCTCTTGGCGGCTTCCCATACCCCAAGCACAGCGCGGATCTA 3291
QY 3447 GCTCTGTACAGACTGGGTGCGGCGAGCGGCTGCTGAGGAGTGTGAGGAGTGTGAGGAGT 3506
Db 3292 GAGCTAGTGCAGAGCATGTGCTGAGGACAACTCTGAGGAGGAGTGTGAGGAGTGTG 3351
QY 3507 GTGCTGTGCTGCTTACACGCGGTGCCATGACGCGAGCTTCTGCGACACTCTTCTCGAGAG 3566
Db 3352 CTGCTTGTGCTACCTTATGCGGGGCGCTGATGGAGCTTCTGCAATGCTTCTTCCAGAG 3411
QY 3567 CTAGACAGCGGCTGCGGAGCTGAGGCTCTACTGGCTACGGGATCTCTCGACACCGACCTC 3626

Db 3412 CTGGATCAGCAGCTGGAGCTCTTGGGTCTCACAGGCTACGGGATCTCGGACACCAACCTG 3471
Qy 3627 GAGGAGATCTTCTGAAGGTGGTGGAGGTGCTCGGACACACAGATATGGAGGTGGC 3586
Db 3472 GAGGAGATCTTCTGAAGGTGGTGGAGG-----TGCACACAGAGAAGGTGGGGACTCT 3525
Qy 3687 AGCTGCGGGCAGCACCTTATGCACAGGCAATTGCTGGCCCTAGACGTAACCCCTCGCGCTCAAG 3746
Db 3526 AGACCGCAGCTGCACCTTCGCACATGC-----ACTCCACAGCCCC 3567
Qy 3747 ATGCCGCCACAGGAGACAGCGCTGGAGAACGGGAAACAGCTGGGTGAGCCCCAGAGACT 3806
Db 3568 ACAGGGCCAGAGGCATCAGTTCTGGAGAAATGGGAGCTGGC----- 3608
Qy 3807 GACCAGGGCTCTGGGCCAGACGCCGTGGCGGGGTACAGGGCTGGGGACGTGACCCGCCAG 3866
Db 3609 --CCACAGGGCTTGGCACCAACGCTGCCCAAGTGCAGAGCTGGACACTTACCTGTGCAA 3666
Qy 3867 CAGCTCCAGGCCCTGTCTTCAAGCGCTTTCGTCTGCCCGCGCAGCGCGCGGGCTG 3926
Db 3667 CAGCTCCGGGCTCTGTCTCCAAAGCGTTTCTGTCTGCTCGCGCAGCCCGCGGGCTG 3726
Qy 3927 TTCGCCAGATCGTGTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATC 3986
Db 3727 TTTGCA CAGGTTGTGTGCTGCCCTCTTTGTGGGCTGGCCCTGTGTTTCAGCCTCAT 3786
Qy 3987 GTGGCTCTTTTCGGGCACTACCGGCTCTCGGGCTAGTCTCCACCATGTAGGTTGCTCAG 4046
Db 3787 GTGGCTCTTTTCGGGCACTACCGGCTCTCGGGCTAGTCTCCACCATGTAGGTTGCTCAG 3846
Qy 4047 GTGTCTCTTTCAGTAGGAGCGCCACAGGGACCCCTGGAGCTGCGCGCTGTCTGAGGG 4106
Db 3847 GTCTCGTTCTTTCAGTAGGATGCCCCCTGGGAGCCCCAACCGATGAAGCTGCTGGAGCT 3906
Qy 4107 CTGCTGAGGAGGAGGACTGGAGGAGCCCCCGAGTGCAGATAGCTCCCAAGG----- 4160
Db 3907 CTGCTAGGGAGGCTGGGCTGCAGGAACCCAGATGTCAGGACAAAGATGCCAGGGATCT 3966
Qy 4161 -----TTCTCGGCACCAAGATTTCTGCTGTAAGTGGCC 4193
Db 3967 GAGTGTACACACTCCCTAGCTTGTACTTTCAGCGTCCCTGAGGTCCTCTGATGTGGCC 4026
Qy 4194 AAGTCTTGGCCAGTGGCAACTGAGCCCGAGGTCTCCATCCCGAGCTGCCAGTGPAGC 4253
Db 4027 AGATCTTGGCCAGTGGCAACTGAGCGCAGAACTCCATCCCGAGCTGCCAATGCAGT 4086
Qy 4254 CAGCCGGTGCCTGGCGCTCTGCTGCCAGCTGCCGCTGCAGCTGGTGTCCCGCTCCG 4313
Db 4087 CAGCCTGGAGCCCGCGCTGTGGCCAGATTGCGCGGCTGGAGCTGGGGGTCCACCAACC 4146
Qy 4314 CCCCAGGCACTGACCGGCTCTGGGAGTGTGTTGAGAACCTGACAGCCCGGAACCTGTCT 4373
Db 4147 CCCCAGGCTGTGGCTGGCTTGGGGAGGTGGTTCAGAACTCACTGGCCGGAATGTGTCT 4206
Qy 4374 GACTTCTGTGTAAGACTTACCCCGCGCTGTGGCGGAGGCTTGAAGACTTAAGAGTGG 4433
Db 4207 GACTTTTGTGTAAGACTTACCCAGCTGTGGCGCGAGGCTTAAGACCAAGAGTGG 4266
Qy 4434 GTGAATGAGTTCAGGTACGAGGCTTCTCGCTGGGGGGCGAGACCCAGGCTGCCCCG 4493
Db 4267 GTGGATGAGTTCAGATATGGGGCTTCTCCTGGAGGCGGAGATCCAGACCTGCCACA 4326
Qy 4494 GGCCAGAGTTGGCCGCTCAGTGGAGGAGTTGTGGCGCTGTGAGTCCCGCTCCCTGGC 4553
Db 4327 GGGCATGAGGTGGTTCGCACATTTGGCAGAGATTGGGCACTGTGAGCGCCCAACCTGGG 4386
Qy 4554 GGGCCCTTCAGGCTGTCTGAAACCTTCAACAGCTGGGCTCACAGCCTGGATGTCTCAG 4613
Db 4387 AATGCGCTAGACCGGTATCTTGAACAACTCACTCAGTGGGCGCTTGGCCCTGTATGCTCG 4446
Qy 4614 GACAGTCTAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTGGCTTTGTCAAC 4673
Db 4447 AACCGGCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTGCCATGGTGGCTTTGTGAC 4506

Qy 4674 CGAGCCAGCAACGCAATCTCTCGTGTCTCACCTGCCCCCAGGGCCCCGCGCACGCCAC 4733
Db 4507 CGAGCCAAACAATGGATCTCTACATGCCCTCTACCACTCTGGTCCAGTCCGCCATGCCAC 4566
Qy 4734 AGCATCACACACACTCAACCAACCCCTTGAACCTCACCAAGAGAGAGTGTCTGAGGTGCA 4793
Db 4567 AGCATCACTACACTCAACCATCTTTGAACCTTGACCAAGAGAGAGCTATCTGAAGCTACA 4626
Qy 4794 CTGATGGCTCTCTCGGTGAGAGTCTCTGCTCCATCTGTGTGTCTTTTGGCATGTCTTTT 4853
Db 4627 CTGATAGCTCTCTCTGTGGATGTCTCTCCATCTGTGTGTCTTTGCGCATGTCTTTT 4686
Qy 4854 GTCCCGGCGAGCTTCACTCTGTCTCATTTGAGGAGGAGTCAACCGAGCAAGACCTG 4913
Db 4687 GTCCCGGCGAGCTTTACCTCTGTCTCTAGAGAAACGATCACAGAGCAAGCATCTG 4746
Qy 4914 CAGCTCATGGGGGCGCTGTCCCCCAGCCCTCTACTGGCTTTGGCAATTTTCTCTGGGACATG 4973
Db 4747 CAGCTGGTCAGCGGCTGTCCCCAAACCTCTATTGGCTTTGGCAATTTCTCTGGGACATG 4806
Qy 4974 TGTAACTACTTGTGTGCAGATGCAATGCTGTGTCTCATCTTTCTGGCTTTTCAGAGAGG 5033
Db 4807 TGTAACTACTTGTGTGCAGTGTGCTAGTGTGTCTTCACTCTTCTAGCTTTTCAGAGAGA 4866
Qy 5034 GCATATGTGGCCCTGTCCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5093
Db 4867 GCATATGTGGCCCGCAGAGAACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4926
Qy 5094 TCGATCACACCGCTCATGTATACCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5153
Db 4927 TCTATCACACCACTCATGTATACCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4986
Qy 5154 GTGGTCTCACTGCTATAAACCTTTTATTTGGCATCAATGGAAGCATGCGCACCTTTTGTG 5213
Db 4987 GTGGTCTCACTGCTATAAACCTTTTATTTGGCATCAATGGAAGCATGCGCACCTTTCTG 5046
Qy 5214 CTTGAGCTCTTCTCTGATCAGAGCTGAGGAGTGCAGCGGATCTTGAAGAGTCTTC 5273
Db 5047 CTAGAACTGCTTTTCAGATCAGAACCTTGAAGAGTGCAGCGGATCTTGAAGAGTCTTC 5106
Qy 5274 CTTATCTTCCCCCAGCTTCTGTCTGGGCGGGGCTCATTCATGATGTTGGGAAACAGGCC 5333
Db 5107 CTTATTTTCCCCCAGCTTGTGCTTGGCCGAGGCTCATTCATGATGTTGGGAAACAGGCC 5166
Qy 5334 ATGGCTGATGCTTTGAGCGCTTGGGAGACAGGAGTTCAGTCAACCTTGCCTGCGCTGGAG 5393
Db 5167 ATGGCAGATGCTTTGAGCGCTTGGGAGACAGCAATTTTCAGTCAACCTTACGCTGGGAC 5226
Qy 5394 GTGGTCGGCAAGAACCTTCTTGGCCATGTTGATACAGGGGCGCTCTTCTCTCTCTCTCT 5453
Db 5227 ATCATTTGGCAAGAACCTTCTTGGCCATGATGCGCCAGGAGCTCTGTCTCTCTCTCT 5286
Qy 5454 CTACTGCTGAGCACCGAAGCCAACTCTCTGCCACAGCCAGGAGTGTCTCTGCCACTC 5513
Db 5287 CTCCTGCTCCNACACCGCAACCGTCTCTCTGCCACATCAAAACCAAGACTGTCTGCCGCC 5346
Qy 5514 CTGGGAGAGAGAGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGTCTCAAGAGCAACC 5573
Db 5347 CTGGGAGAGAGAGATGAGGATGTGGCTCAAGAGCGTGCAGCGGGTGCACCAAGGGGCTACC 5406
Qy 5574 CAGGGGATGTGTGTGTGTGAGGAACTTGAACCAAGGTATACCGTGGGAGAGGATGCCA 5633
Db 5407 CAGGGGATGTGTGTGTGTGCTCTAGGAGCTTGAACCAAGGTATACCGTGGGAGAGAAACCA 5466
Qy 5634 GCTTGTGAACCGCTTGTGCTTGGGATTCCTCTGTGAGTGTGTGTGTGTGTGTGTGTGTG 5693
Db 5467 GCTGTGATGCTG 5526
Qy 5694 AATGAGAGAGAGAGAGAGCTCCAGTTCGATGTGTGAACGGGGAGACATTTGGCCAGGAG 5753
Db 5527 AACGGGCGAGGGAAGACATCCACCTTCGCGATGTGTGACAGGGGACACACTGCCCCAGCT 5586

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:41:30 ; Search time 2868 Seconds
(without alignments)
12453.647 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctcaggggcgccgcgtcccc.....cgtgctggtgaaacaaaaa 6804

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6804	100.0	6804	8	ABX14666 Human CDN
2	6762.4	99.4	6791	6	AAL44693 Human tra
3	6717.8	98.7	6768	6	AAS19207 DNA encod
4	6717.8	98.7	6768	10	Add37428 Human tra
5	6710	98.6	7795	6	ABL57810 Human ABC
6	6478.4	95.2	6522	4	AAS08706 Human PD-
7	6398.4	94.0	6432	10	Add37430 Human tra
8	6036.8	88.7	6327	8	ABX72257 Human NOV
9	5933.2	87.2	6696	8	ABX95283 CDNA enco
10	5833.2	85.7	6651	8	ABX95284 CDNA enco
11	5529	81.3	5811	4	AAD05626 Human sec
12	5529	81.3	5811	8	ADA40555 Human sec
13	5529	81.3	5811	10	ADA56706 Gene enco
14	5488.2	80.2	5669	4	AAS08707 Human PD-
15	4843.4	71.2	5859	8	ABX95285 Human ABC
16	4381.6	64.4	4653	8	ABX14667 Human CDN
17	3902.4	57.4	6633	8	ABX14665 Mouse CDN
18	3885.8	57.1	6607	4	ABF54812 Nucleotid
19	3474.4	50.7	3635	4	AAF56389 ABC trans
20	3303	48.5	5762	4	AAF54792 Nucleotid
21	3129.8	46.0	3437	5	AAS73965 DNA encod

22	3039.8	44.7	4413	5	AAS83715	Aas83715 DNA encod
23	2861.8	42.1	2911	3	AAZ94751	Aaz94751 Human ATP
24	1808.6	26.6	1879	11	ADM03146	Adm03146 Human CDN
25	1740.6	25.6	6786	9	ACC84923	Acc84923 Human ABC
26	1739	25.6	10412	9	ACC84922	Acc84922 Human ABC
27	1739	25.6	10412	11	ADP65742	Adp65742 Human ATP
28	1739	25.6	10412	11	ADP64975	Adp64975 Human ATP
29	1739	25.6	10412	12	ADM41251	Adm41251 Human ATP
30	1737.4	25.5	7860	4	AAF92835	Aaf92835 Human ABC
31	1737.4	25.5	7860	5	AAF83826	Aaf83826 Human ABC
32	1737.4	25.5	10365	10	ADD94068	Add94068 Human ATP
33	1737.4	25.5	10381	10	ADD94069	Add94069 Human ATP
34	1737.4	25.5	10423	10	ADD94070	Add94070 Human ATP
35	1737.4	25.5	10442	4	AAF24680	Aaf24680 Nucleotid
36	1737.4	25.5	10442	4	AAF24702	Aaf24702 Nucleotid
37	1735.8	25.5	9741	6	ABL58146	AbL58146 Human ABC
38	1735.8	25.5	9870	6	ABN99319	Abn99319 Polymorph
39	1735.8	25.5	9870	6	ABN99316	Abn99316 Polymorph
40	1735.8	25.5	9870	6	ABN99315	Abn99315 Polymorph
41	1735.8	25.5	9870	6	ABN99308	Abn99308 Polymorph
42	1735.8	25.5	9870	6	ABN99318	Abn99318 Polymorph
43	1735.8	25.5	10474	4	AAF24685	Aaf24685 Nucleotid
44	1735.8	25.5	10474	4	AAF24686	Aaf24686 Nucleotid
45	1735.8	25.5	10474	4	AAF24708	Aaf24708 Nucleotid

ALIGNMENTS

RESULT 1
ABX14666
ID ABX14666 standard; CDNA; 6804 BP.
AC ABX14666;
XX
DT 10-MAR-2003 (first entry)
XX
DE Human cDNA encoding ATP-binding cassette transporter-like protein.
XX
KW Human; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; ss; gene;
KW axonopathy; autoimmune disease; inflammatory disease; multiple sclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 210..6650
FT /*tag= a
FT /product= "ABCL"
FT sig_peptide 210..347
FT /*tag= b
FT mat_peptide 348..6647
FT /*tag= c
FT /label= Mature_ABCL
XX
US2002127647-A1.
XX
PD 12-SEP-2002.
XX
PF 28-NOV-2001; 2001US-00995542.
XX
PR 28-NOV-2000; 2000US-0253520P.
XX
PA (SHUT/) SHUTTER J.
PA (ULIA/) ULIAS L.

Db 1441 AGCTGAGGGGACCCCTCAGAGGACAGCCCTGGTGTGCGGGCCCTGCACTGCTCGCGG 1500
Qy 1501 AACATCGATTCTGGGCGGCGCTGCTTCTTGGGACCTTAGAGACTCTTTCAGACCCACAG 1560
Db 1501 AACATCGATTCTGGGCGGCGCTGCTTCTTGGGACCTTAGAGACTCTTTCAGACCCACAG 1560
Qy 1561 AGCACCCAAACCCAGACCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
Db 1561 AGCACCCAAACCCAGACCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
Qy 1621 AGCTGGTCAAGAGGACCAATAGATCAGGACAGAGTCTTGGGACCTTGGGCGGCGGCGGCGG 1680
Db 1621 AGCTGGTCAAGAGGACCAATAGATCAGGACAGAGTCTTGGGACCTTGGGCGGCGGCGGCGG 1680
Qy 1681 ACCCCCTGACCGACTGCGCTACGTGTGGGGCGGCTTGGTACCTGCAAGACCTGGTGG 1740
Db 1681 ACCCCCTGACCGACTGCGCTACGTGTGGGGCGGCTTGGTACCTGCAAGACCTGGTGG 1740
Qy 1741 AGCGTGCAGCGTCCGCGTGTCTAGCGGGGCCAACCCTGGGCGGCGGCGGCGGCGGCGG 1800
Db 1741 AGCGTGCAGCGTCCGCGTGTCTAGCGGGGCCAACCCTGGGCGGCGGCGGCGGCGGCGG 1800
Qy 1801 AGATGCCCTATCCGCTGCTATGTGGACGAGCTGTCTCGCTGTGCTGAGCGGCTCGCTGC 1860
Db 1801 AGATGCCCTATCCGCTGCTATGTGGACGAGCTGTCTCGCTGTGCTGAGCGGCTCGCTGC 1860
Qy 1861 CGCTCTTCTGACCGTGGCGCTGATCTACTCGTGACACTGACAGTGAAGCGGCTGGTGC 1920
Db 1861 CGCTCTTCTGACCGTGGCGCTGATCTACTCGTGACACTGACAGTGAAGCGGCTGGTGC 1920
Qy 1921 GGGAGAAGAGACGCGGCTGGGGGACACATGCGCGGCAATGGGGCTCAGCGCGGCGTGC 1980
Db 1921 GGGAGAAGAGACGCGGCTGGGGGACACATGCGCGGCAATGGGGCTCAGCGCGGCGTGC 1980
Qy 1981 TCTGCTAGGCTGGTTCCTCAGTGCCTCGGGCGGCTTCTGCTCAGCGCGGCTGGTGG 2040
Db 1981 TCTGCTAGGCTGGTTCCTCAGTGCCTCGGGCGGCTTCTGCTCAGCGCGGCTGGTGG 2040
Qy 2041 TTCTGGTCTCAAGCTGGGGACATCTCCCTACAGCACCCGCGGCGTGGTCTTCTGT 2100
Db 2041 TTCTGGTCTCAAGCTGGGGACATCTCCCTACAGCACCCGCGGCGTGGTCTTCTGT 2100
Qy 2101 TCTTGGAGCCTTCCGCGTGGCCAGCGTGAACCGAGCTTCTGCTCAGCGGCTTCTTCT 2160
Db 2101 TCTTGGAGCCTTCCGCGTGGCCAGCGTGAACCGAGCTTCTGCTCAGCGGCTTCTTCT 2160
Qy 2161 CCGCGCCAACTGGCTGGCGCTGGCGGCGGCTTCTGCTCAGCGGCTTCTTCTTCTTCT 2220
Db 2161 CCGCGCCAACTGGCTGGCGCTGGCGGCGGCTTCTGCTCAGCGGCTTCTTCTTCTTCT 2220
Qy 2221 ACCTGCTGTGTGGGCGGACCGGCTGCGCGGCGGCTGGCGGCGGCGGCGGCGGCGG 2280
Db 2221 ACCTGCTGTGTGGGCGGACCGGCTGCGCGGCGGCTGGCGGCGGCGGCGGCGGCGG 2280
Qy 2281 TGCTGTCCCGTGGCTTCCGCTTCCGCTGCGAGAGCTTGGCTTCTGCTGAGGAGCAGG 2340
Db 2281 TGCTGTCCCGTGGCTTCCGCTTCCGCTGCGAGAGCTTGGCTTCTGCTGAGGAGCAGG 2340
Qy 2341 GCGAGGCGCGCAGTGGGACAACTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2400
Db 2341 GCGAGGCGCGCAGTGGGACAACTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2400
Qy 2401 CCCAGGCTCTGGGCTTCTGCTGTGGAGCGGCGGCTTCTGCGGCTGCGGCGGCGGCGG 2460
Db 2401 CCCAGGCTCTGGGCTTCTGCTGTGGAGCGGCGGCTTCTGCGGCTGCGGCGGCGGCGG 2460
Qy 2461 TGGAGCTGTGTGCCAGGACGATACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Db 2461 TGGAGCTGTGTGCCAGGACGATACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Qy 2521 GGAGCTACTGTGGGACCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580
Db 2521 GGAGCTACTGTGGGACCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580

Qy 2581 ACCAAAGGTGCTGTAGAGAGGCAACCGCCCGGCTTGTGCTTCTGCGGCTTATCCGTTGCA 2640
Db 2581 ACCAAAGGTGCTGTAGAGAGGCAACCGCCCGGCTTGTGCTTCTGCGGCTTATCCGTTGCA 2640
Qy 2641 GCCTGGAGAGGCTTCTTCTGGAAGCCGAGCCGAGCCCTGCGGGGGCTCAGCCTGGACT 2700
Db 2641 GCCTGGAGAGGCTTCTTCTGGAAGCCGAGCCGAGCCCTGCGGGGGCTCAGCCTGGACT 2700
Qy 2701 TCTACAGGGCCACATCACCCCTTCTTGGGCGGCAACAGGGGCGGCAAGACACACCC 2760
Db 2701 TCTACAGGGCCACATCACCCCTTCTTGGGCGGCAACAGGGGCGGCAAGACACACCC 2760
Qy 2761 TGTCCATCTTGTAGTGGCTTCTTCCACCCAGTGGTGGTCTGCTTCTTCTTCTGCGGCGG 2820
Db 2761 TGTCCATCTTGTAGTGGCTTCTTCCACCCAGTGGTGGTCTGCTTCTTCTTCTGCGGCGG 2820
Qy 2821 AGCTCCGCTCAGACATGCGGCGGCAATCCGCGGCGGCAACAGGGGCGGCAAGACACAC 2880
Db 2821 AGCTCCGCTCAGACATGCGGCGGCAATCCGCGGCGGCAACAGGGGCGGCAAGACAC 2880
Qy 2881 TGCTGTTTGCATGCTGACCGTGGAGCAGCTGCTGGTCTTATGGGCGGCTGAAGGGTCT 2940
Db 2881 TGCTGTTTGCATGCTGACCGTGGAGCAGCTGCTGGTCTTATGGGCGGCTGAAGGGTCT 2940
Qy 2941 TGAGTGGCGCTGTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3000
Db 2941 TGAGTGGCGCTGTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3000
Qy 3001 CCAAGAGAGTGTGAGACTCGGCACTCTCTGTGGGATGCAACGGAAGCTGCTCGTGG 3060
Db 3001 CCAAGAGAGTGTGAGACTCGGCACTCTCTGTGGGATGCAACGGAAGCTGCTCGTGG 3060
Qy 3061 CCATTTGCTTGTGGGCGGCTCCCAAGTGTGTTATCTTGGAGAGCTTACGCTGCGGCTG 3120
Db 3061 CCATTTGCTTGTGGGCGGCTCCCAAGTGTGTTATCTTGGAGAGCTTACGCTGCGGCTG 3120
Qy 3121 ATCTGCTTCCCGCGGCGGCTATTTGGGAGCTGCTGCTCAAAATACCGAAGGCTCGCAGC 3180
Db 3121 ATCTGCTTCCCGCGGCGGCTATTTGGGAGCTGCTGCTCAAAATACCGAAGGCTCGCAGC 3180
Qy 3181 TGATCTCTCCACCCACCTGATGAGGAGAGCTGCTGGGAGACCGTGTGGCTGG 3240
Db 3181 TGATCTCTCCACCCACCTGATGAGGAGAGCTGCTGGGAGACCGTGTGGCTGG 3240
Qy 3241 TGGCAGGTGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
Db 3241 TGGCAGGTGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
Qy 3301 CCGGCTACTACTGACCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3360
Db 3301 CCGGCTACTACTGACCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3360
Qy 3361 ACATGTACATGGAGGAGTGTGAGCACCGAGGAGGAAAGAAATGGCAGCAGGCGCA 3420
Db 3361 ACATGTACATGGAGGAGTGTGAGCACCGAGGAGGAAAGAAATGGCAGCAGGCGCA 3420
Qy 3421 GCAGAGTGGGAGCTTCTTCTGAGTGGGCTTGGTACAGCACTGGGTCGCGGCGGCGGCGG 3480
Db 3421 GCAGAGTGGGAGCTTCTTCTGAGTGGGCTTGGTACAGCACTGGGTCGCGGCGGCGGCGG 3480
Qy 3481 TGGTGGAGGAGTCTCCACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
Db 3481 TGGTGGAGGAGTCTCCACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
Qy 3541 GCAGCTTCCGACACTCTTCCGAGAGCTAGACACCGCGGCTGGCGGAGCTGAGGCTCACTG 3600
Db 3541 GCAGCTTCCGACACTCTTCCGAGAGCTAGACACCGCGGCTGGCGGAGCTGAGGCTCACTG 3600
Qy 3601 GCTACGGGAGTCTCCGACACCGAGCTTCTTCTGAGAGTCTTCTTCTGAGGAGTGTG 3660
Db 3601 GCTACGGGAGTCTCCGACACCGAGCTTCTTCTGAGAGTCTTCTTCTGAGGAGTGTG 3660

Qy	3661	CTGGGACACAGATATATGAGGATGGCAGCTGCGGGCAGCACCTATATGACAGGCAATTGCTG	3720
Db	3661	CTGGGGACACAGATATATGAGGATGGCAGCTGCGGGCAGCACCTATATGACAGGCAATTGCTG	3720
Qy	3721	GCCTAGACGTAAACCTTCGCGCTCAAGATGCGCGCACAGGAGACAGCGCTGGAGAACGGGG	3780
Db	3721	GCCTAGACGTAAACCTTCGCGCTCAAGATGCGCGCACAGGAGACAGCGCTGGAGAACGGGG	3780
Qy	3781	AACAGCTGGGTGAGCCCGCCAGAGACTGACCAAGGGCTCTGGGCGACAGCCGCTGGGCGGGG	3840
Db	3781	AACAGCTGGGTGAGCCCGCCAGAGACTGACCAAGGGCTCTGGGCGACAGCCGCTGGGCGGGG	3840
Qy	3841	TACAGGCTGGGCACTGACCGCGCAGCAGCTCAAGGCGCTTCTCAAGCGCTTTCTGC	3900
Db	3841	TACAGGCTGGGCACTGACCGCGCAGCAGCTCAAGGCGCTTCTCAAGCGCTTTCTGC	3900
Qy	3901	TTGCCCGCCGACGCGCGCGGCGCTGTTCGCCACAGATCGTGCCTCGCCCTCTTTGTGG	3960
Db	3901	TTGCCCGCCGACGCGCGCGGCGCTGTTCGCCACAGATCGTGCCTCGCCCTCTTTGTGG	3960
Qy	3961	GCCTGGCCCTCGTGTTCAGCGCTCATCTGCTCTCTTTTCGGGCACTACCCGGCTCTGGGC	4020
Db	3961	GCCTGGCCCTCGTGTTCAGCGCTCATCTGCTCTCTTTTCGGGCACTACCCGGCTCTGGGC	4020
Qy	4021	TCAGTCCACCATGTACGGTGTCTCAGGTGTCTTCTTCAGTGAGGAGCGCCAGGGGACC	4080
Db	4021	TCAGTCCACCATGTACGGTGTCTCAGGTGTCTTCTTCAGTGAGGAGCGCCAGGGGACC	4080
Qy	4081	CTGACGTGCGCGGCTGCTCGAGCGCTGTCTGACGAGGAGGAGCTGAGGAGGCGCCACG	4140
Db	4081	CTGACGTGCGCGGCTGCTCGAGCGCTGTCTGACGAGGAGGAGCTGAGGAGGCGCCACG	4140
Qy	4141	TGCAGATAGTCTCCACAGGTTCTCGGCAACAGAGTTCTCTGTAAGTGCGCCAGGTCT	4200
Db	4141	TGCAGATAGTCTCCACAGGTTCTCGGCAACAGAGTTCTCTGTAAGTGCGCCAGGTCT	4200
Qy	4201	TGGCCAGTGGAACCTGAGACCCAGAGTCTCCATCCCGAGCTGCGAGTGAGGAGGCGCG	4260
Db	4201	TGGCCAGTGGAACCTGAGACCCAGAGTCTCCATCCCGAGCTGCGAGTGAGGAGGCGCG	4260
Qy	4261	GTGCCCGGCGCTGCTGCCCGACTGCCCGGTGCGAGCTGTGTGCTCCCGCCAGG	4320
Db	4261	GTGCCCGGCGCTGCTGCCCGACTGCCCGGTGCGAGCTGTGTGCTCCCGCCAGG	4320
Qy	4321	CAGTGAACCGGCTCTGGGAGTGTTCAGAACCTGACAGCGCGGAACTGTCTGACTTCC	4380
Db	4321	CAGTGAACCGGCTCTGGGAGTGTTCAGAACCTGACAGCGCGGAACTGTCTGACTTCC	4380
Qy	4381	TGGTCAAGACTACCCGCGCTGTGGCGCAGAGGCTGAAAGCTAAGAACTGGGTGAATG	4440
Db	4381	TGGTCAAGACTACCCGCGCTGTGGCGCAGAGGCTGAAAGCTAAGAACTGGGTGAATG	4440
Qy	4441	AGGTCAAGTACGAGGCTTCTCGCTGGGGGCGCGAGACCCAGGCGCTCCCTCGGGCCAAAG	4500
Db	4441	AGGTCAAGTACGAGGCTTCTCGCTGGGGGCGCGAGACCCAGGCGCTCCCTCGGGCCAAAG	4500
Qy	4501	AGTTGGGCGGCTCAGTGAGAGTGTGGGGCTGTGAGTCCCTCGCTGGGCGGGCCC	4560
Db	4501	AGTTGGGCGGCTCAGTGAGAGTGTGGGGCTGTGAGTCCCTCGCTGGGCGGGCCC	4560
Qy	4561	TCGACCGTGTCTCAAAAACCTCACAGCTGGGCTCACAGCTGGATGTCTCAGGACAGTC	4620
Db	4561	TCGACCGTGTCTCAAAAACCTCACAGCTGGGCTCACAGCTGGATGTCTCAGGACAGTC	4620
Qy	4621	TCAGATCTGTTCAACAACAAGGCTGGCACTCCATGTGTGGCTTTGTCAACCGAGCCA	4680
Db	4621	TCAGATCTGTTCAACAACAAGGCTGGCACTCCATGTGTGGCTTTGTCAACCGAGCCA	4680
Qy	4681	GCAACGCAATCTCGTGTCTACCTGCGCCCGAGCGCGCGCGCCAGCCACAGCATCA	4740
Db	4681	GCAACGCAATCTCGTGTCTACCTGCGCCCGAGCGCGCGCGCCAGCCACAGCATCA	4740
Qy	4741	CCACACTCAACCCCTTGAACCTCACCAAGGAGCAGCTGTCTGAGGCTGCATGATGG	4800

Db	4741	CCACACTCAACCAACCCCTTGAACCTCACCAAGGAGCAGCTGTCTGAGGCTGCATGATGG	4800
Qy	4801	CTTCTCGGTGAGCGTCTCGTCTCCATCTGTGTGTCTTTGTCATGTCTTTGTCCCGG	4860
Db	4801	CTTCTCGGTGAGCGTCTCGTCTCCATCTGTGTGTCTTTGTCATGTCTTTGTCCCGG	4860
Qy	4861	CCAGCTTCACTCTTGTCTCAITGAGGAGCGAGTCAACCGAGCAAGCACTTGCAGTCA	4920
Db	4861	CCAGCTTCACTCTTGTCTCAITGAGGAGCGAGTCAACCGAGCAAGCACTTGCAGTCA	4920
Qy	4921	TGGGGGCGCTGTCCCCACCCCTCTACTTGGCTTGGCAACTTCTCTGGGACATGTAACT	4980
Db	4921	TGGGGGCGCTGTCCCCACCCCTCTACTTGGCTTGGCAACTTCTCTGGGACATGTAACT	4980
Qy	4981	ACTTGTGCCAGCATGATGCTGTGCTCATCTTCTTGGCTTCCAGCAGAGGGCATATG	5040
Db	4981	ACTTGTGCCAGCATGATGCTGTGCTCATCTTCTTGGCTTCCAGCAGAGGGCATATG	5040
Qy	5041	TGGCCCTCGCAACCTCGCTCTCTCTGTGTGTCTACTACTGTATGGCTGTGATCA	5100
Db	5041	TGGCCCTCGCAACCTCGCTCTCTCTGTGTGTCTACTACTGTATGGCTGTGATCA	5100
Qy	5101	CACCGCTCATGTACCCAGCGCTCTTCTTCTTCCGTGCCAGCAGACCTATGTGTGC	5160
Db	5101	CACCGCTCATGTACCCAGCGCTCTTCTTCTTCCGTGCCAGCAGACCTATGTGTGC	5160
Qy	5161	TCACTTGCATATAACCTCTTTATTTGGCATATGGAAGCATGCGCACTTTGTGCTTGAGC	5220
Db	5161	TCACTTGCATATAACCTCTTTATTTGGCATCAATGGAAGCATGCGCACTTTGTGCTTGAGC	5220
Qy	5221	TCTTCTCTGATCAGAGCTGCAGAGGTGAGCGGATCTTTGAAACAGGTCTCTCTATCT	5280
Db	5221	TCTTCTCTGATCAGAGCTGCAGAGGTGAGCGGATCTTTGAAACAGGTCTCTCTATCT	5280
Qy	5281	TCCCCCACTTCTGCTTGGGCGCGGCTCATTTGACATGTGTGCGGAACCCAGGCCATGCTG	5340
Db	5281	TCCCCCACTTCTGCTTGGGCGCGGCTCATTTGACATGTGTGCGGAACCCAGGCCATGCTG	5340
Qy	5341	ATGCCCTTGTAGCGCTTGGGAGACAGGAGTTCAGTCAACCCCTGCGTGGGAGGTGTG	5400
Db	5341	ATGCCCTTGTAGCGCTTGGGAGACAGGAGTTCAGTCAACCCCTGCGTGGGAGGTGTG	5400
Qy	5401	GCAAGAACCTTGTGGCATGTGATACAGGGGCGCTCTTCTTCTTCTTCACTACTG	5460
Db	5401	GCAAGAACCTTGTGGCATGTGATACAGGGGCGCTCTTCTTCTTCTTCACTACTG	5460
Qy	5461	TGCAGCACCGAAGCCAACTCTGCCACAGCCAGGGGTGAGGTCTCTGCCACTCTGGGAG	5520
Db	5461	TGCAGCACCGAAGCCAACTCTGCCACAGCCAGGGGTGAGGTCTCTGCCACTCTGGGAG	5520
Qy	5521	AGGAGACGAGGATGTAGCCGTGAACGGGAGCGGTGTCCAAAGAGCCACCAGGGGG	5580
Db	5521	AGGAGACGAGGATGTAGCCGTGAACGGGAGCGGTGTCCAAAGAGCCACCAGGGGG	5580
Qy	5581	ATCTGTGTGTGCTGAGGAACTTCACCAAGTATACCGTGGGAGAGATGCCAGCTTGT	5640
Db	5581	ATCTGTGTGTGCTGAGGAACTTCACCAAGTATACCGTGGGAGAGATGCCAGCTTGT	5640
Qy	5641	ACCGCTTGTGCTTGGGATTTCCCGCTTGGTGTGTTTGGGCTGTGGGTGTGAATGGAG	5700
Db	5641	ACCGCTTGTGCTTGGGATTTCCCGCTTGGTGTGTTTGGGCTGTGGGTGTGAATGGAG	5700
Qy	5701	CAGGAGACGTCACATTTTCGATGTGTGACGGGGAGACATTTGGCCAGCAGGGCGAGG	5760
Db	5701	CAGGAGACGTCACATTTTCGATGTGTGACGGGGAGACATTTGGCCAGCAGGGCGAGG	5760
Qy	5761	CTGTGTGGCAGCCACAGCGTGGCGGAAACCCAGTGTCTGCGACCTCAGCATGGAT	5820
Db	5761	CTGTGTGGCAGCCACAGCGTGGCGGAAACCCAGTGTCTGCGACCTCAGCATGGAT	5820
Qy	5821	ACTGCCCTCAATCCGATGCTTGTGAGTGTCTGACGGGCGCGAGCACCTGAGGTGC	5880
Db	5821	ACTGCCCTCAATCCGATGCTTGTGAGTGTCTGACGGGCGCGAGCACCTGAGGTGC	5880

Db 5821 ACTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACCGGCGCGGAGCACCTTGAGCTGC 5880
Qy 5881 TTGCGCGCTGCGCGCTGCTCCCGAGGCGCCAGTTTGCACAGACCCCTGGCTCAGGCGCTGG 5940
Db 5881 TTGCGCGCTGCGCGCTGCTCCCGAGGCGCCAGTTTGCACAGACCCCTGGCTCAGGCGCTGG 5940
Qy 5941 CGCGTCTGGACTCTCATGTATGACGAGACCGCGCTGAGGACCTTACAGCGAGGGAACA 6000
Db 5941 CGCGTCTGGACTCTCATGTATGACGAGACCGCGCTGAGGACCTTACAGCGAGGGAACA 6000
Qy 6001 AAGCAAGCTGCGAGCGCGCTGCGCTGCTGTTGGGACCCAGCGCTGTTCTTGGAGC 6060
Db 6001 AAGCAAGCTGCGAGCGCGCTGCGCTGCTGTTGGGACCCAGCGCTGTTCTTGGAGC 6060
Qy 6061 AGCCACACAGGATGAGACCCAGCGCGCGCTTCTTTGGAAAGCCCTTTTGGCGG 6120
Db 6061 AGCCACACAGGATGAGACCCAGCGCGCGCTTCTTTGGAAAGCCCTTTTGGCGG 6120
Qy 6121 TGGTGGGAGGCGCGCTTCACTGATGCTCACTCCCATAGCATGAGGAGTGTGAAGCGC 6180
Db 6121 TGGTGGGAGGCGCGCTTCACTGATGCTCACTCCCATAGCATGAGGAGTGTGAAGCGC 6180
Qy 6181 TCTGCTCGCGCTAGCCATCATGTTGATGAGGCGGTTCCGCTGCTGGGAGCGCGCAAC 6240
Db 6181 TCTGCTCGCGCTAGCCATCATGTTGATGAGGCGGTTCCGCTGCTGGGAGCGCGCAAC 6240
Qy 6241 ATCTCAAGGCGAGATTTCGCGCGGCTCACACATGACCTCGGCTGCGCGCGCAAGGT 6300
Db 6241 ATCTCAAGGCGAGATTTCGCGCGGCTCACACATGACCTCGGCTGCGCGCGCAAGGT 6300
Qy 6301 CCCAGCGGAGCGCGCTTCTGCGCGCGGAGTTCCTTGGGTCGAGCTGCGCGAGGCAC 6360
Db 6301 CCCAGCGGAGCGCGCTTCTGCGCGCGGAGTTCCTTGGGTCGAGCTGCGCGAGGCAC 6360
Qy 6361 ATGAGGCGCGCTGCGCTTCCAGCTGCGCGCGGAGGCGCTGCGCGCGCGCT 6420
Db 6361 ATGAGGCGCGCTGCGCTTCCAGCTGCGCGCGGAGGCGCTGCGCGCGCGCT 6420
Qy 6421 TTGAGAGCTGCGCGTGACGCGCGGAGAGACGCGGCTGAGGACTTTTCCGTGAGCCAGA 6480
Db 6421 TTGAGAGCTGCGCGTGACGCGCGGAGAGACGCGGCTGAGGACTTTTCCGTGAGCCAGA 6480
Qy 6481 CGATGCTGAGAGGATTTCTTTGATCTTCAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAG 6540
Db 6481 CGATGCTGAGAGGATTTCTTTGATCTTCAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAG 6540
Qy 6541 AAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6600
Db 6541 AAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6600
Qy 6601 GCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6660
Db 6601 GCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6660
Qy 6661 CCGTGGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6720
Db 6661 CCGTGGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6720
Qy 6721 CTGAGCTCAGGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6780
Db 6721 CTGAGCTCAGGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6780
Qy 6781 AAGCGCTGCTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6804
Db 6781 AAGCGCTGCTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6804

RESULT 2

AAL44693

ID AAL44693 standard; cDNA; 6791 BP.

XX

AC AAL44693;

XX

03-MAY-2002 (first entry)

Human transporter and ion channel TRICH-27 cDNA.

Human; transporter and ion channel; TRICH; transport disorder;
neurological disorder; muscle disorder; immunological disorder;
cell proliferative disorder; neuroprotective; nootropic;
cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
gene therapy; gene; 88.

Homo sapiens.

WO2000204520-A2.

17-JAN-2002.

05-JUL-2001; 2001WO-US021448.

07-JUL-2000; 2000US-0216547P.

14-JUL-2000; 2000US-0218232P.

21-JUL-2000; 2000US-0220112P.

28-JUL-2000; 2000US-0221839P.

(INCY-) INCYTE GENOMICS INC.

Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;

Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J;

Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Walia NK;

Au-Young J, Walsh RT, Runkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;

Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;

Kearney L, Thangavelu K, Das D, Policky JL;

WPI: 2002-205969/26.

P-PSDB; AAO14210.

New human transporters and ion channel polypeptides for diagnosing,
treating or preventing transport, neurological, muscle, immunological and
cell proliferative disorders.

Claim 5; Page 224-225; 230pp; English.

The present invention provides the protein and coding sequences of a
number of human transporter and ion channel proteins, designated TRICH-1-
TRICH-32. The sequences can be used in the treatment of transport,
neurological, muscle, immunological and cell proliferative disorders. The
present sequence is a coding sequence of the invention

Sequence 6791 BP; 1137 A; 2202 C; 2184 G; 1268 T; 0 U; 0 Other;

Query Match 99.4%; Score 6762.4; DB 6; Length 6791;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 6780; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 12 CGCGCTCCCTGCTGCTGCTGGGCGAGGAGGCGGCAAGAGCTGCGGAGCCCTGGAA 71

Db 1 CGCGCTCCCTGCTGCTGCTGGGCGAGGAGGCGGCAAGAGCTGCGGAGCCCTGGAA 60

Qy 72 GAGCTTCCAGAACCCCTGCGCTGGGATAGGAATCAGGTTTCAAGAGGCGGAGGAG 131

Db 61 GAGCTTCCAGAACCCCTGCGCTGGGATAGGAATCAGGTTTCAAGAGGCGGAGGAG 119

Qy 132 TTGCGCGGAGCGGAGCGGAGCGGAGCTTTCAGCCGCGGAGGAGGAGGAGGAGGAG 191

Db 120 TTGCGCGGAGCGGAGCGGAGCGGAGCTTTCAGCCGCGGAGGAGGAGGAGGAG 179

Qy 192 CCCCAGGCGGAGCTTCAACATGCGCTTTCGAGACAGAGTGTGCTGCTCTGGAAGAT 251

Db 180 CCCCAGGCGGAGCTTCAACATGCGCTTTCGAGACAGAGTGTGCTGCTGGAAGAT 239

Qy 252 TTCATGTATCGCGGAGGAGCGGAGCTTTCAGCTTCTGCTGGAATTTGCTGTGCGCTCTTC 311

Db 240 TTCATGTATCGCGGAGGAGCGGAGCTTTCAGCTTCTGCTGGAATTTGCTGTGCGCTCTTC 299

Qy	312	CTCTTTCTTCATCTGGTGGCTGTTTCGCCACTCCCAACCGCCCTGGAGCACCATTGAATGC	371
Db	300	CTCTTTCTTCATCTGGTGGCTGTTTCGCCACTCCCAACCGCCCTGGAGCACCATTGAATGC	359
Qy	372	CACCTTCCAAACAAAGCACTGCCATCGCGGGGCAACGTCGCCCTGGCTCCAGGGTCTCATC	431
Db	360	CACCTTCCAAACAAAGCACTGCCATCGCGGGGCAACGTCGCCCTGGCTCCAGGGTCTCATC	419
Qy	432	TGTAATGTGAACAACACTGCTTTCGCCAGCTGACACCGGGGAGAGCCCGGGCGCTG	491
Db	420	TGTAATGTGAACAACACTGCTTTCGCCAGCTGACACCGGGGAGAGCCCGGGCGCTG	479
Qy	492	AGCAACTTCAACGACTCCCTGGTCTCCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGGA	551
Db	480	AGCAACTTCAACGACTCCCTGGTCTCCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGGA	539
Qy	552	GGGGCCAGTGCCACAGGACGCTGGCTGGCTAGGGAAAGCTGATCGCCACGCTGAGGGCT	611
Db	540	GGGGCCAGTGCCACAGGACGCTGGCTGGCTAGGGAAAGCTGATCGCCACGCTGAGGGCT	599
Qy	612	GCAGCAGCAGCGCCAGCCTCAACAAACCAAGCAGTCTCACTGGBAAACCAACCAATGCTG	671
Db	600	GCAGCAGCAGCGCCAGCCTCAACAAACCAAGCAGTCTCACTGGBAAACCAACCAATGCTG	659
Qy	672	GATGTCGCGAGCTGCTGAGCTCACTGCTGGCAACGAATCCCTGGGGTTGGCACTGGGC	731
Db	660	GATGTCGCGAGCTGCTGAGCTCACTGCTGGCAACGAATCCCTGGGGTTGGCACTGGGC	719
Qy	732	CAAGCCAGGAGCCCTTTGCACAGCTTGTGTGGAGCCGCTGAGGACCTGGCCCAAGGAGCTC	791
Db	720	CAAGCCAGGAGCCCTTTGCACAGCTTGTGTGGAGCCGCTGAGGACCTGGCCCAAGGAGCTC	779
Qy	792	CTGGCGCTGGCGAGCTGGTGGAGCTTTCGGGCACTGCTGCAGAGACCCCGAGGAGCAGC	851
Db	780	CTGGCGCTGGCGAGCTGGTGGAGCTTTCGGGCACTGCTGCAGAGACCCCGAGGAGCAGC	839
Qy	852	GGCCCTCTGAGTTGCTGTCAGAGGCCCTCTGCACTGTCAAGGGACCTAGCAGCACAGTG	911
Db	840	GGCCCTCTGAGTTGCTGTCAGAGGCCCTCTGCACTGTCAAGGGACCTAGCAGCACAGTG	899
Qy	912	GGCCCTCTCTCAACTGCTACGAGGCTAGTGAACCTGATGAGCTGTGGGGCAGGAGCCA	971
Db	900	GGCCCTCTCTCAACTGCTACGAGGCTAGTGAACCTGATGAGCTGTGGGGCAGGAGCCA	959
Qy	972	GAATCCGCCCTGCCAGACAGGAGCTGAGCCCGCCCTGCTCGAGAGCTGATGGAGCCCTG	1031
Db	960	GAATCCGCCCTGCCAGACAGGAGCTGAGCCCGCCCTGCTCGAGAGCTGATGGAGCCCTG	1019
Qy	1032	GACAGCACCCGCTGCTCCGCTGCTCTGGAGACGCTGAAAGCCTCTGATCCTCGGGAAG	1091
Db	1020	GACAGCACCCGCTGCTCCGCTGCTCTGGAGACGCTGAAAGCCTCTGATCCTCGGGAAG	1079
Qy	1092	CTACTCTTTGACACAGATACACCTTTTACCCGGAAGCTCATGGCCCAAGGTGAACCGGACC	1151
Db	1080	CTACTCTTTGACACAGATACACCTTTTACCCGGAAGCTCATGGCCCAAGGTGAACCGGACC	1139
Qy	1152	TTGAGAGCTCACCTCTGAGGGATGTCCGGAGGTGTGGAGATGCTGGAGACCCCGG	1211
Db	1140	TTGAGAGAGTTCACCTCTGAGGGATGTCCGGAGGTGTGGAGATGCTGGAGACCCCGG	1199
Qy	1212	ATCTTTACCTTCATGAACGACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGCAGATG	1271
Db	1200	ATCTTTACCTTCATGAACGACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGCAGATG	1259
Qy	1272	CAGGATGAAGGAAGAGGAGCCAGACCTGGAGCCCGGGAACCAATGGAGGCGCTTCGGA	1331
Db	1260	CAGGATGAAGGAAGAGGAGCCAGACCTGGAGCCCGGGAACCAATGGAGGCGCTTCGGA	1319
Qy	1332	TCCTTTCTGGACCTCTGGAGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGG	1391
Db	1320	TCCTTTCTGGACCTCTGGAGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGG	1379
Qy	1392	CACCTGGTGGGACGCTGGGCGGAGTGACGAGTGCTGCTCTTGGACAAGCTGGAGGCG	1451

Db	1380	CACCTGGTGGGACGCTGGGCGGAGTGACGGAGTGCTGCTTGGACAAGCTGGAGCG	1439
Qy	1452	GCACCTCTAGAGGACAGCCCTGGTGTGCGGGGCCCTGCAACTGCTCGGGGAACATCGATTTC	1511
Db	1440	GCACCTCTAGAGGACAGCCCTGGTGTGCGGGGCCCTGCAACTGCTCGGGGAACATCGATTTC	1499
Qy	1512	TGGGCGGGGTGCTCTTCTTGGGACCTGAGGACTCTTTCAGACCCCAAGAGCAGCCCAACC	1571
Db	1500	TGGGCGGGGTGCTCTTCTTGGGACCTGAGGACTCTTTCAGACCCCAAGAGCAGCCCAACC	1559
Qy	1572	CCAGAACTTGGGCGCCCGGCCCAACGTCGCATCAAAATCCGATGACATTTGACGTGGTCAAG	1631
Db	1560	CCAGAACTTGGGCGCCCGGCCCAACGTCGCATCAAAATCCGATGACATTTGACGTGGTCAAG	1619
Qy	1632	AGAACCAATAAGATCAGGAGACAGTGTGGGACCTTGGGCCAGCCGCGGACCCCTTGACC	1691
Db	1620	AGAACCAATAAGATCAGGAGACAGTGTGGGACCTTGGGACCCCTTGGCCAGCCGCGGACCC	1679
Qy	1692	GACCTGGCTACGTCGTGGGCGGCTTCTGTACTCTGACAGACCTGGTGGAGCGTGCAGCC	1751
Db	1680	GACCTGGCTACGTCGTGGGCGGCTTCTGTACTCTGACAGACCTGGTGGAGCGTGCAGCC	1739
Qy	1752	GTCCGGTGTCTCAGCGGCGCCAAACCCCGGGCGGCTCTTACCTGCAGCAGATGCCCTAT	1811
Db	1740	GTCCGGTGTCTCAGCGGCGCCAAACCCCGGGCGGCTCTTACCTGCAGCAGATGCCCTAT	1799
Qy	1812	CGTGTATGTGGAAGACGTCGTGGTGTCTGAGCGGTGCTGCGCTCTTCCTG	1871
Db	1800	CGTGTATGTGGAAGACGTCGTGGTGTCTGAGCGGTGCTGCGCTCTTCCTG	1859
Qy	1872	ACGCTGGCTGGATCTTACTCCGTGACACTGACAGTGAAGCCGCTGGTGGGGAAGAGAG	1931
Db	1860	ACGCTGGCTGGATCTTACTCCGTGACACTGACAGTGAAGCCGCTGGTGGGGAAGAGAG	1919
Qy	1932	ACCGGCTGCGGACACCATGCGCGCATGGGGCTCAGCGCGCGGTGCTCTGGCTAGGC	1991
Db	1920	ACCGGCTGCGGACACCATGCGCGCATGGGGCTCAGCGCGCGGTGCTCTGGCTAGGC	1979
Qy	1992	TGTTTCTCAGCTGCTCGGGCCCTTCTGCTCAGCGCGCGGTGCTGGTCTGGTGTCTC	2051
Db	1980	TGTTTCTCAGCTGCTCGGGCCCTTCTGCTCAGCGCGCGGTGCTGGTCTGGTGTCTC	2039
Qy	2052	AAGCTGGGGACATCTCCCTACAGCACCCGGGGTGGTCTCTGTTCTTGGCAGCC	2111
Db	2040	AAGCTGGGGACATCTCCCTACAGCACCCGGGGTGGTCTCTGTTCTTGGCAGCC	2099
Qy	2112	TTCCGGTGGCCACGGTGAACCCAGAGCTTCTGCTCAGCGCCTTCTTCTCCCGCGCAAC	2171
Db	2100	TTCCGGTGGCCACGGTGAACCCAGAGCTTCTGCTCAGCGCCTTCTTCTCCCGCGCAAC	2159
Qy	2172	CTGGCTGGGCGCTGGCGGCGCTTGGCTTCTTCTCCTCTACCTGCCCTACGCTGTGTGT	2231
Db	2160	CTGGCTGGGCGCTGGCGGCGCTTGGCTTCTTCTCCTCTACCTGCCCTACGCTGTGTGT	2219
Qy	2232	GTGGCTTGGGGACCGGCTGCGCGGGTGGCGGTGGCGCGGAGCTGCTGTGCGCC	2291
Db	2220	GTGGCTTGGGGACCGGCTGCGCGGGTGGCGGTGGCGCGGAGCTGCTGTGCGCC	2279
Qy	2292	GTGGCTTGGGCTTGGCTGCGAGAGCTTGGCTCTGCTGGAGGAGCAGGGCGGCGCG	2351
Db	2280	GTGGCTTGGGCTTGGCTGCGAGAGCTTGGCTCTGCTGGAGGAGCAGGGCGGCGCG	2339
Qy	2352	CAGTGGCAAACTGGGCGCCCGGCTTACGGCAGAGCTTCTTACGCTGGCCCGAGGCTCT	2411
Db	2340	CAGTGGCAAACTGGGCGCCCGGCTTACGGCAGAGCTTCTTACGCTGGCCCGAGGCTCT	2399
Qy	2412	GGCTTCTGCTGTCAGCGCGGCTCTACGGCTCGCCACCTGGTACCTTGGAGCTGTG	2471
Db	2400	GGCTTCTGCTGTCAGCGCGGCTCTACGGCTCGCCACCTGGTACCTTGGAGCTGTG	2459
Qy	2472	TGCCAGGCGCAGTACGGGATCCCTGAACCATGGAAATTTCTCTTTCGGAGGAGCTACTGG	2531

Db 2460 TCGCCAGGCGAGTACGGGATCCTGAA CCAATGAAATTTTCTTTTCGGAGGAGCTACTGG 2519
Qy 2532 TCGGAGCTTCGCCCCCAAGAGTCCAGCCCTTCGCCACCCCGCTGAGACCAAGGTG 2591
Db 2520 TCGGAGCTTCGCCCCCAAGAGTCCAGCCCTTCGCCACCCCGCTGAGACCAAGGTG 2579
Qy 2592 CTGGTAGAAGAGGACCGCCCGGCTGAGTCTCGGGTATCCGTTCCGACGCTGAGAG 2651
Db 2580 CTGGTAGAAGAGGACCGCCCGGCTGAGTCTCGGGTATCCGTTTCGACGCTGAGAG 2639
Qy 2652 CGCTTTCTGGAAGCCCGCAGCAGCCCTTCGGGGGCTCAGCTTCGATCTTACCAAGGC 2711
Db 2640 CGCTTTCTGGAAGCCCGCAGCAGCCCTTCGGGGGCTCAGCTTCGATCTTACCAAGGC 2699
Qy 2712 CACATCACCGCTTCCTGCGGCACAA CGGGGCGGCAAGACCA CACCCCTGTCATCTTG 2771
Db 2700 CACATCACCGCTTCCTGCGGCACAA CGGGGCGGCAAGACCA CACCCCTGTCATCTTG 2759
Qy 2772 AGTGGCTCTTCCACCCAGTGGTGGCTTCGCTTCATCTGGGCGACAGTCGGCTCC 2831
Db 2760 AGTGGCTCTTCCACCCAGTGGTGGCTTCGCTTCATCTGGGCGACAGTCGGCTCC 2819
Qy 2832 AGCATGGCGCCATCCGGCCCACTTGGGCGTCTGCTTCAGTCAACAGTGTGTTGAC 2891
Db 2820 AGCATGGCGCCATCCGGCCCACTTGGGCGTCTGCTTCAGTCAACAGTGTGTTGAC 2879
Qy 2892 ATGCTGACCGTGGAGCAGCAGCTCTGCTTCTATGGGCGGCTGAAGGCTCTGAGTCCGCT 2951
Db 2880 ATGCTGACCGTGGAGCAGCAGCTCTGCTTCTATGGGCGGCTGAAGGCTCTGAGTCCGCT 2939
Qy 2952 GTAGTGGGCGCGAGCAGGACCGTCTGCTGAGAGATGTGGGCTGGTCTCAAGCAGAGT 3011
Db 2940 GTAGTGGGCGCGAGCAGGACCGTCTGCTGAGAGATGTGGGCTGGTCTCAAGCAGAGT 2999
Qy 3012 GTGCAGACTCGCCACTCTCTGTTGGATGAA CCGAAGCTGCTCGTGGCCATTCGCTTT 3071
Db 3000 GTGCAGACTCGCCACTCTCTGTTGGATGAA CCGAAGCTGCTCGTGGCCATTCGCTTT 3059
Qy 3072 GTGGGCGGCTCCAAAGTTGTTATCTTGA CAGAGCTTACGGCTGGGCTGGATCTCTCTCC 3131
Db 3060 GTGGGCGGCTCCAAAGTTGTTATCTTGA CAGAGCTTACGGCTGGGCTGGATCTCTCTCC 3119
Qy 3132 CGCCGCGGATTTGGGAGTGTCTCAAA TACGAGAGGTGCGACGCTGATCTCTCC 3191
Db 3120 CGCCGCGGATTTGGGAGTGTCTCAAA TACGAGAGGTGCGACGCTGATCTCTCC 3179
Qy 3192 ACCCACCCTGGATGAGGAGAGCTGCTGGAGACCGTGGCTGTGGTGGCAGGTGGC 3251
Db 3180 ACCCACCCTGGATGAGGAGAGCTGCTGGAGACCGTGGGAGACCGTGGTGGCAGGTGGC 3239
Qy 3252 CGCTTGTGCTGTGGTCTCCCACTCTTCTG CCGCGCTCACCCTGGGCTCCGGCTACTAC 3311
Db 3240 CGCTTGTGCTGTGGTCTCCCACTCTTCTG CCGCGCTCACCCTGGGCTCCGGCTACTAC 3299
Qy 3312 CTGACGCTGGTGAAGGCGCGCTGCCCTGAC CCAATGAGAGGCTGACATGACATG 3371
Db 3300 CTGACGCTGGTGAAGGCGCGCTGCCCTGAC CCAATGAGAGGCTGACATGACATG 3359
Qy 3372 GAGGCACTGTGGACACGAGCAGAAAGAGATG CAGCCAGGCGAGCAGAGTCCGC 3431
Db 3360 GAGGCACTGTGGACACGAGCAGAAAGAGATG CAGCCAGGCGAGCAGAGTCCGC 3419
Qy 3432 ACTCTCAGCTGTGGGCTCTGATCAGCACTG GGGTCCCGGGCACGGCTGGTGGAGAG 3491
Db 3420 ACTCTCAGCTGTGGGCTCTGATCAGCACTG GGGTCCCGGGCACGGCTGGTGGAGAG 3479
Qy 3492 CTGCCACAGAGCTGGTGTGGTCTGCCCTTAC CAGGGTGGCCATGACGGCAGCTTCGCC 3551
Db 3480 CTGCCACAGAGCTGGTGTGGTCTGCCCTTAC CAGGGTGGCCATGACGGCAGCTTCGCC 3539
Qy 3552 ACACCTTCCGAGAGCTAGACCGGGCTGGG GAGCTGAGGCTCAGCTGGCTACGGGATC 3611
Db 3540 ACACCTTCCGAGAGCTAGACCGGGCTGGG GAGCTGAGGCTCAGCTGGCTACGGGATC 3599

Qy 3612 TCCGACACGAGCTCGAGGAGATCTTCTGA AGTGGTGGAGAGTGTGCTCGGACACA 3671
Db 3600 TCCGACACGAGCTCGAGGAGATCTTCTGA AGTGGTGGAGAGTGTGCTCGGACACA 3659
Qy 3672 GATATGGAGGATGGCAGCTGGGGCAGACA CTTATGACAGGATTTGCTGGCCCTAGACGTA 3731
Db 3660 GATATGGAGGATGGCAGCTGGGGCAGACA CTTATGACAGGATTTGCTGGCCCTAGACGTA 3719
Qy 3732 ACCCTGCGGCTCAAGATGCCCA CAGGAGACAGCGCTGGGAAACGGGAAACAGCTGG 3791
Db 3720 ACCCTA CCGCTCAAGATGCCCA CAGGAGACAGCGCTGGGAAACGGGAAACAGCTGG 3779
Qy 3792 TCAGCCCCAGAGACTGACAGGCTCTGGG CCAAGCGCTGGGCGGGGTACAGGGCTGG 3851
Db 3780 TCAGCCCCAGAGACTGACAGGCTCTGGG CCAAGCGCTGGGCGGGGTACAGGGCTGG 3839
Qy 3852 GCATGACCCCGCAGCAGCTCAGGCCCTGCT TCTCAAGCGCTTCTGCTTGGCCCGCGC 3911
Db 3840 GCATGACCCCGCAGCAGCTCAGGCCCTGCT TCTCAAGCGCTTCTGCTTGGCCCGCGC 3899
Qy 3912 AGCCGCGCGGCTGTTCCGCCAGATCGTGC TCTGCTGCTTGTGCGGCTGSCCTC 3971
Db 3900 AGCCGCGCGGCTGTTCCGCCAGATCGTGC TCTGCTGCTTGTGCGGCTGSCCTC 3959
Qy 3972 GTGTTGAGCTCATGCTGCTCTTTCGGG CACACTACCGGCTCTCGGCTCAGTCCAC 4031
Db 3960 GTGTTGAGCTCATGCTGCTCTTTCGGG CACACTACCGGCTCTCGGCTCAGTCCAC 4019
Qy 4032 ATGTACGCTGTCTAGGTGCTCTTCTTCA GTAGAGACGCCCAAGGGAACCTTGGAGCTGCC 4091
Db 4020 ATGTACGCTGTCTAGGTGCTCTTCTTCA GTAGAGACGCCCAAGGGAACCTTGGAGCTGCC 4079
Qy 4092 CGGCTGCTTCAGGCGCTGCTGCAGGAGCAGG ACTGGAGGAGCCCGAGTCAGCATAGC 4151
Db 4080 CGGCTGCTTCAGGCGCTGCTGCAGGAGCAGG ACTGGAGGAGCCCGAGTCAGCATAGC 4139
Qy 4152 TCCACAGCTTCTCGGCACCAAGAGTTCTC TCTGAAGTGGGCAAGGTCTTGGCCAGTGGC 4211
Db 4140 TCCACAGCTTCTCGGCACCAAGAGTTCTC TCTGAAGTGGGCAAGGTCTTGGCCAGTGGC 4199
Qy 4212 AACTGGACCCAGAGTCTCATTCGCCAGCTG CCAAGTGTAGCAGCCCGGTGCGCGGCGC 4271
Db 4200 AACTGGACCCAGAGTCTCATTCGCCAGCTG CCAAGTGTAGCAGCCCGGTGCGCGGCGC 4259
Qy 4272 CTGCTGCCGACTGCCCGGCTGAGCTGGTGC CCCCCTCCGCCAGGAGTGGCCGCGC 4331
Db 4260 CTGCTGCCGACTGCCCGGCTGAGCTGGTGC CCCCCTCCGCCAGGAGTGGCCGCGC 4319
Qy 4332 TCTGGGGAAGTGGTTGAGAACCTTGACAGG CCGGAAACCTGTCTGACTTCTGCTCAAGACC 4391
Db 4320 TCTGGGGAAGTGGTTGAGAACCTTGACAGG CCGGAAACCTGTCTGACTTCTGCTCAAGACC 4379
Qy 4392 TACCCGCGCTGTGTGGCGCAGGGCTGAAGA CTAAAGTGGGTGAATGAGGTCAAGTAC 4451
Db 4380 TACCCGCGCTGTGTGGCGCAGGGCTGAAGA CTAAAGTGGGTGAATGAGGTCAAGTAC 4439
Qy 4452 GGAGGCTTCTCGCTGGGGGCGGAGACCCAGG CCGTCCCTCGGGCCAAAGAGTGGGGCGC 4511
Db 4440 GGAGGCTTCTCGCTGGGGGCGGAGACCCAGG CCGTCCCTCGGGCCAAAGAGTGGGGCGC 4499
Qy 4512 TCAGTGGAGGAGTGTGTGGCGCTGCTGAGTC CCGCTTCTGCTGGGGGCGCTCGACCGTGT 4571
Db 4500 TCAGTGGAGGAGTGTGTGGCGCTGCTGAGTC CCGCTTCTGCTGGGGGCGCTCGACCGTGT 4559
Qy 4572 CTGAAAACTCTCAGCGCTGGGCTCAGAGCTG GATGCTCAGGACAGTCTCAAGATCTGG 4631
Db 4560 CTGAAAACTCTCAGCGCTGGGCTCAGAGCTG GATGCTCAGGACAGTCTCAAGATCTGG 4619
Qy 4632 TTCAACAAAGAGCTGGCACTCCATGGTGGCTT TGTCAACCGGAGCCAGCAACGCAATC 4691
Db 4620 TTCAACAAAGAGCTGGCACTCCATGGTGGCTT TGTCAACCGGAGCCAGCAACGCAATC 4679

QY 4692 CTCGTCCTCCTCCTGCCCCCAGGGCCCGCGCCGACAGCCACACAGCATCACACACTCAAC 4751
DB 4680 CTCGTCCTCCTCCTGCCCCCAGGGCCCGCGCCGACAGCCACACAGCATCACACACTCAAC 4739
QY 4752 CACCCCTTGAACCTTCAACCAAGGAGCAGCTGCTGAGGCTGCACCTGATGGCTCTCTCGGTG 4811
DB 4740 CACCCCTTGAACCTTCAACCAAGGAGCAGCTGCTGAGGCTGCACCTGATGGCTCTCTCGGTG 4799
QY 4812 GAGTCCTCTGTCCTCATCTGCTGTGGTCTTTGGCCATGTCCTTTGTCGGGCCAGCTTCACCT 4871
DB 4800 GAGTCCTCTGTCCTCATCTGCTGTGGTCTTTGGCCATGTCCTTTGTCGGGCCAGCTTCACCT 4859
QY 4872 CTTGTCTCTCATGAGGAGCAGTCAACCCGAGCACAAGCACCTGAGCTCATGGGGGGCTG 4931
DB 4860 CTTGTCTCTCATGAGGAGCAGTCAACCCGAGCACAAGCACCTGAGCTCATGGGGGGCTG 4919
QY 4932 TCCCCCAACCTCTACTGGCTTGGCACTTTCTCTGGGACATGTGTAACTACTTTGGTGCCA 4991
DB 4920 TCCCCCAACCTCTACTGGCTTGGCACTTTCTCTGGGACATGTGTAACTACTTTGGTGCCA 4979
QY 4992 GCATGATCTGTGTCCTATCTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCTGCCC 5051
DB 4980 GCATGATCTGTGTCCTATCTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCTGCCC 5039
QY 5052 AACCTGCTCTCTCTGCTGTGCTACTACTGTATGTGCTGTGCTGCATCACACGCTCATG 5111
DB 5040 AACCTGCTCTCTCTGCTGTGCTACTACTGTATGTGCTGTGCTGCATCACACGCTCATG 5099
QY 5112 TACCAGCCTCTCTCTCTCTCTCGTGCCAGCAGCACCTATGTGGTGTCTCAGCTGCATTA 5171
DB 5100 TACCAGCCTCTCTCTCTCTCGTGCCAGCAGCACCTATGTGGTGTCTCAGCTGCATTA 5159
QY 5172 AACCTCTTTATGGATCAATGAGAGCATGGCCACCTTTGTGCTGTGAGCTCTTCTGTAT 5231
DB 5160 AACCTCTTTATGGATCAATGAGAGCATGGCCACCTTTGTGCTGTGAGCTCTTCTGTAT 5219
QY 5232 CAGAAGCTGAGAGGTGAGCCGATCTTGAACAGGTCTTCCCTATCTTCCCCCACTTC 5291
DB 5220 CAGAAGCTGAGAGGTGAGCCGATCTTGAACAGGTCTTCCCTATCTTCCCCCACTTC 5279
QY 5292 TGCTTGGGCGGGGGCTCATTTGACATGGTGGGAACAGAGCCATGGCTGATGCTTTGAG 5351
DB 5280 TGCTTGGGCGGGGGCTCATTTGACATGGTGGGAACAGAGCCATGGCTGATGCTTTGAG 5339
QY 5352 CGCTTGGGAGACAGGAGTTCAGTCAACCCCTGGCTGGAGGTGTCGGCAAGACCTC 5411
DB 5340 CGCTTGGGAGACAGGAGTTCAGTCAACCCCTGGCTGGAGGTGTCGGCAAGACCTC 5399
QY 5412 TTGGCCATGTGTATACAGGGGCCCTCTCTCTCTCTTCACTACTGCTGCAGCACCGA 5471
DB 5400 TTGGCCATGTGTATACAGGGGCCCTCTCTCTCTCTTCACTACTGCTGCAGCACCGA 5459
QY 5472 AGCCAACTCTTGCACAGCCAGGGTGAAGTCTCTGCCACTCTCTGGAGAGGAGGACGAG 5531
DB 5460 AGCCAACTCTTGCACAGCCAGGGTGAAGTCTCTGCCACTCTCTGGAGAGGAGGACGAG 5519
QY 5532 GATGTAGCCCGTGAAACGGGAGCGGGTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTG 5591
DB 5520 GATGTAGCCCGTGAAACGGGAGCGGGTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTG 5579
QY 5592 CTGAGGAACCTTGACCAAGGTATACCGTGGGAGAGGATGCAGCTGTTGACCCGCTTGTGC 5651
DB 5580 CTGAGGAACCTTGACCAAGGTATACCGTGGGAGAGGATGCAGCTGTTGACCCGCTTGTGC 5639
QY 5652 CTGGGGAATCCCCCTGTGAGTGTGTTTGGCTGCTGGGTGTGAATGGAGCAGGGAAGACG 5711
DB 5640 CTGGGGAATCCCCCTGTGAGTGTGTTTGGCTGCTGGGTGTGAATGGAGCAGGGAAGACG 5699
QY 5712 TCACAGTTTCGATGTGTGACGGGGGACACATTTGGCCAGAGGGGCGAGGCTGTCTGGCA 5771
DB 5700 TCACAGTTTCGATGTGTGACGGGGGACACATTTGGCCAGAGGGGCGAGGCTGTCTGGCA 5759
QY 5772 GGCCACAGCTGGGCCCGGGAACCCAGTGTCTGCTGCGCACCTCAGCATGGGATACTGCCCTCAA 5831

DB 5760 GGCCACAGCGTGGCCCGGGNAACCCAGTGTGTCGCACCTCAGCATGGATCTGCCCTCAA 5819
QY 5832 TCCGATGCCATCTTTGAGTGTGTGACGGGCGGAGCACTGAGAGTGTGTTGGCGGCTG 5891
DB 5820 TCCGATGCCATCTTTGAGTGTGTGACGGGCGGAGCACTGAGAGTGTGTTGGCGGCTG 5879
QY 5892 CGCGGTGTCCCGAGGCCCCAGGTTGGCCAGACCGCTGGCTCAGGCGTGGCGGCTCTGGGA 5951
DB 5880 CGCGGTGTCCCGAGGCCCCAGGTTGGCCAGACCGCTGGCTCAGGCGCTCTGGGA 5939
QY 5952 CTCTCATGTGTACGACAGCCGGCTGAGGCACTTACAGCGGAGGGAACAAACGCAAGCTG 6011
DB 5940 CTCTCATGTGTACGACAGCCGGCTGAGGCACTTACAGCGGAGGGAACAAACGCAAGCTG 5999
QY 6012 GCGACGGCCCTGCGCTGTGTTGGGACCCAGGCGTGTGTTCTGAGCAGAGCCGACACA 6071
DB 6000 GCGACGGCCCTGCGCTGTGTTGGGACCCAGGCGTGTGTTCTGAGCAGAGCCGACACA 6059
QY 6072 GCGATGGAACCCAGCGCGCGCTTCTTCTTGGAAACAGCCTTTTGGCCGTTGGCGGGAG 6131
DB 6060 GCGATGGAACCCAGCGCGCGCTTCTTCTTGGAAACAGCCTTTTGGCCGTTGGCGGGAG 6119
QY 6132 GCGCGTTCACTGTCTCACTCCCATAGCATGGAGAGTGTGAAGCGCTCTCTCTCGCGC 6191
DB 6120 GCGCGTTCACTGTCTCACTCCCATAGCATGGAGAGTGTGAAGCGCTCTCTCTCGCGC 6179
QY 6192 CTAGCCATCATGTGTAATGGGGGTTCCGCTGCTGGGAGCCGCAACATCTCAAGGGC 6251
DB 6180 CTAGCCATCATGTGTAATGGGGGTTCCGCTGCTGGGAGCCGCAACATCTCAAGGGC 6239
QY 6252 AGATTGCGCGGGTCAACACACTGACCTTCCGCTGCGGTCGCCGCAAGGTCCCAAGCCGCA 6311
DB 6240 AGATTGCGCGGGTCAACACACTGACCTTCCGCTGCGGTCGCCGCAAGGTCCCAAGCCGCA 6299
QY 6312 GCGGCTTCTGTCGCGCCGAGTTCCCTGCTGCGTGGAGCTGCGCGAGGCAACATGAGGCGCG 6371
DB 6300 GCGGCTTCTGTCGCGCCGAGTTCCCTGTCGGGCGGAGTGGCGAGGCAACATGAGGCGCG 6359
QY 6372 CTGCGCTTCACTGTCGCGCGGAGGGGCTGCGGCTGCGGCGGCTCTTTTGGAGAGCTG 6431
DB 6360 CTGCGCTTCACTGTCGCGCGGAGGGCTGCGGCTGCGGCGGCTCTTTTGGAGAGCTG 6419
QY 6432 GCGGTGCAAGGCGCAGACGCGCTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAG 6491
DB 6420 GCGGTGCAAGGCGCAGACGCGCTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAG 6479
QY 6492 GAGGTATTCTTGTACTTCTTCAAGGACCAAGGGAAGGAGGACCCGAAAGAGCAGAG 6551
DB 6480 GAGGTATTCTTGTACTTCTTCAAGGACCAAGGGAAGGAGGACCCGAAAGAGCAGAG 6539
QY 6552 GAGGCAAGGAGTGGGAGTGGACCCCGCGCAGGCTGAGGACCCCAAAACGCGTACGCGAG 6611
DB 6540 GAGGCAAGGAGTGGGAGTGGACCCCGCGCAGGCTGAGGACCCCAAAACGCGTACGCGAG 6599
QY 6612 TTTCTCGATGACCCCTAGCATGCGGAGACTGTCTCTGAGCCTTCCCTCCCTCGGGGGCC 6671
DB 6600 TTTCTCGATGACCCCTAGCATGCGGAGACTGTCTCTGAGCCTTCCCTCCCTCGGGGGCC 6559
QY 6672 GCGGGGAGGCGCTTGGGAATGGCAAGGCAAGGTAGAGTGTCTAGGAGCCCTGGACTCAGG 6731
DB 6660 GCGGGGAGGCGCTTGGGAATGGCAAGGCAAGGTAGAGTGTCTAGGAGCCCTGGACTCAGG 6719
QY 6732 CTGGCAGAGGGCTGTGCTCCCTGGGAAATAAAGAGAGGCTGGAGAGGCGCTGCTG 6791
DB 6720 CTGGCAGAGGGCTGTGCTCCCTGGGAAATAAAGAGAGGCTGGAGAGGCGCTGCTG 6779
QY 6792 GTGAAAAAAA 6803
DB 6780 GTGAAAAAAA 6791

Db	1075	AGCTCATGGCCCAAGGTGAACCGGACCTTCGAGGAGCTCACCTCTGCTGAGGGATGTCTCGGG	1134
Qy	1186	AGGTGTGGGAGATGCTCTGGGACCCCGGATCTTACCTTCATGAACGACAGTTTCAATGTGG	1245
Db	1135	AGGTGTGGGAGATGCTCTGGGACCCCGGATCTTACCTTCATGAACGACAGTTTCAATGTGG	1194
Qy	1246	CCATGCTGCAGCGGCTCTCTGCAGATGCAGATGAAGGAAGAGCAGCCAGACCTTGGAG	1305
Db	1195	CCATGCTGCAGCGGCTCTCTGCAGATGCAGATGAAGGAAGAGCAGCCAGACCTTGGAG	1254
Qy	1306	GCCGGGACCAATGGAGGCCCTTGGAATCCTTTCTGGACCTTGGAGCGGTGGCTACAGCT	1365
Db	1255	GCCGGGACCAATGGAGGCCCTTGGAATCCTTTCTGGACCTTGGAGCGGTGGCTACAGCT	1314
Qy	1366	GGCAGGACGACACGCTGATGTGGGCACCTGTGGGCAGCGCTGGGCCGAGTGACCGAGT	1425
Db	1315	GGCAGGACGACACGCTGATGTGGGCACCTGTGTGGGCACTGTGGGCCGAGTGACCGAGT	1374
Qy	1426	GCCTGTCTTTGGACAAGCTGGAGCGGCACCTCAGAGGCAGCCCTGGTGTCTGGGCCGAGTGACCGAGT	1485
Db	1375	GCCTGTCTTTGGACAAGCTGGAGCGGCACCTCAGAGGCAGCCCTGGTGTCTGGGCCGAGTGACCGAGT	1434
Qy	1486	TGCAACTGCTCGCGGAAACATCGATTTGGGCCGCGCTGCTTCTTTGGGACCTTGAGGACT	1545
Db	1435	TGCAACTGCTCGCGGAAACATCGATTTGGGCCGCGCTGCTTCTTTGGGACCTTGAGGACT	1494
Qy	1546	CTTCAGACCCACAGAGACCCCAACCCAGACCTGGGCCCGGCGGACCTGGGCATCAAAA	1605
Db	1495	CTTCAGACCCACAGAGACCCCAACCCAGACCTGGGCCCGGCGGACCTGGGCATCAAAA	1554
Qy	1606	TCCGATATGGACATTGAGCTGTACGAGGACCAATAAGATCAGGGACAGGTTTGGGACC	1665
Db	1555	TCCGATATGGACATTGAGCTGTACGAGGACCAATAAGATCAGGGACAGGTTTGGGACC	1614
Qy	1666	CTGCCCCAGCCGCGACCCCTGACGACCTGCGCTACGTGTGGGCGGCTTCTGTGTACC	1725
Db	1615	CTGCCCCAGCCGCGACCCCTGACGACCTGCGCTACGTGTGGGCGGCTTCTGTGTACC	1674
Qy	1726	TGCNAGACTGTGTGGAGCGTGACGCGTCCGCTGCTCAGGGGGCCCAACCCCGGGCCG	1785
Db	1675	TGCNAGACTGTGTGGAGCGTGACGCGTCCGCTGCTCAGGGGGCCCAACCCCGGGCCG	1734
Qy	1786	GCCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGGACGACGTTCCTCGCTGTGC	1845
Db	1735	GCCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGGACGACGTTCCTCGCTGTGC	1794
Qy	1846	TGAGCCGGTCCGTGCCGCTCTTCTGACGCTGGCGTGGATCTACTCCGTGACACTGACAG	1905
Db	1795	TGAGCCGGTCCGTGCCGCTCTTCTGACGCTGGCGTGGATCTACTCCGTGACACTGACAG	1854
Qy	1906	TGAAGCCGTGTGTGGGAGAAAGAGACGCGGCTCGGGGACACCAATCGCGCCATCGGGC	1965
Db	1855	TGAAGCCGTGTGTGGGAGAAAGAGACGCGGCTCGGGGACACCAATCGCGCCATCGGGC	1914
Qy	1966	TCAGCCGCGGGTGTCTCTGGCTAGGCTGTTCTCTACGCTCGGCGCCTTCTGTCTCA	2025
Db	1915	TCAGCCGCGGGTGTCTCTGGCTAGGCTGTTCTCTACGCTCGGCGCCTTCTGTCTCA	1974
Qy	2026	GCGCCGCGCTGCTGGTGTCTGATGCTCAAGCTGGGGGACATCTCTCCCTCAGGCCACCCGG	2085
Db	1975	GCGCCGCGCTGCTGGTGTCTGATGCTCAAGCTGGGGGACATCTCTCCCTCAGGCCACCCGG	2034
Qy	2086	GCGTGTCTTCTGTTCTTGGACGCTTTCGCGTGGGCAACCGGTGACCCAGAGTTTCTGTC	2145
Db	2035	GCGTGTCTTCTGTTCTTGGACGCTTTCGCGTGGGCAACCGGTGACCCAGAGTTTCTGTC	2094
Qy	2146	TCAGCGCTTCTTCTTCTCCGCGCAACCTGCTGGGCTCGGGGCGCTGGCCTACTTCT	2205
Db	2095	TCAGCGCTTCTTCTTCTCCGCGCAACCTGCTGGGCTCGGGGCGCTGGCCTACTTCT	2154
Qy	2206	CCCTCTACCTGCCCTACGCTGTGTGTGGCTTGGCGGGACCGGCTGCCCGCGGGTGCC	2265

2155	CCCTCTACTGCGCCTACGTGCTGTGTGTGGCTGTGGCGGGACA	CGGGCTGCCCGGGGTGGCC	2214	
2266	GCGTGGCCGCGAGACCTGCTGTGCGCCGTGGCCCTTCGGCTTC	CGGTGCGAGAGCCTGGCTC	2325	
2215	GCGTGGCCGCGAGACCTGCTGTGCGCCGTGGCCCTTCGGCTTC	CGGTGCGAGAGCCTGGCTC	2274	
2326	TGCTGGAGGAGCAGGGCGCGCAGCTGTGCACAA	CGTGGGCAACC	CGGCTTACGGCAG	2385
2275	TGCTGGAGGAGCAGGGCGCGCAGCTGTGCACAA	CGTGGGCAACC	CGGCTTACGGCAG	2334
2386	ACGTCTTACAGCTGGGCCCAAGTCTCTGGCCCTTCGTCTGCT	TGGACGCGGGCTCTACGGCC	2445	
2335	ACGTCTTACAGCTGGGCCCAAGTCTCTGGCCCTTCGTCTGCT	TGGACGCGGGCTCTACGGCC	2394	
2446	TGCGCCACCTGTACTGTGAAGCTGTGTGCCAGGCGACGTAC	GGGATCCCTGAACCATGGA	2505	
2395	TGCGCCACCTGTACTGTGAAGCTGTGTGCCAGGCGACGTAC	GGGATCCCTGAACCATGGA	2454	
2506	ATTTTCTCTTTTCGGAGGAGCTACTGCTGCGGACCTCGGCC	CCCCCAAGAGTCCAGCCCTT	2565	
2455	ATTTTCTCTTTTCGGAGGAGCTACTGCTGCGGACCTCGGCC	CCCCCAAGAGTCCAGCCCTT	2514	
2566	GCCGCAACCCGCTGGACCCAAAGAGTGTGTGTAGAGAGGCA	CCGCCCGGCTGTAGTCTGTG	2625	
2515	GCCGCAACCCGCTGGACCCAAAGTGTGTGTAGAGAGGCA	CCGCCCGGCTGTAGTCTGTG	2574	
2626	GGCTATCCGTTTCGAGAGCTGGAGAGCGCTTTCTGTGAAG	CGCCGAGCGACCTCGGG	2685	
2575	GGCTATCCGTTTCGAGAGCTGGAGAGCGCTTTCTGTGAAG	CGCCGAGCGACCTCGGG	2634	
2686	GGCTCAGCCTGGACTTCTACAGGGGCGACATCACCGCTTCT	TGGGCGCAACCGGGGCGG	2745	
2635	GGCTCAGCCTGGACTTCTACAGGGGCGACATCACCGCTTCT	TGGGCGCAACCGGGGCGG	2694	
2746	GCAAGACAACACCTGTGTCATCTTTAGTGGGCTTTCCCA	CCACCAAGTGGTGGCTGTGCT	2805	
2695	GCAAGACAACACCTGTGTCATCTTTAGTGGGCTTTCCCA	CCACCAAGTGGTGGCTGTGCT	2754	
2806	TCATCTGGGGCCACGACGTCGCTCCAGCATGGCGCGCCAT	CGGCGCCACCTGGGCGTCT	2865	
2755	TCATCTGGGGCCACGACGTCGCTCCAGCATGGCGCGCCAT	CGGCGCCACCTGGGCGTCT	2814	
2866	GTCTCAGTACAACCTGTCTGTTTGACATGCTGACCGTGG	AGCAGCAGCTCTGGTTCTATG	2925	
2815	GTCTCAGTACAACCTGTCTGTTTGACATGCTGACCGTGG	AGCAGCAGCTCTGGTTCTATG	2874	
2926	GGCGGCTGAAGGGTCTGAGTCCGCTGTAGTGGGCGCCCG	AGAGAGACCGTCTGTGCGAGG	2985	
2875	GGCGGCTGAAGGGTCTGAGTCCGCTGTAGTGGGCGCCCG	AGAGAGACCGTCTGTGCGAGG	2934	
2986	ATGTGGGGCTGTCTCCAGCAGAGTGTGCAGACTCGCCAC	CTCTCTGGTGGGATGCAAC	3045	
2935	ATGTGGGGCTGTCTCCAGCAGAGTGTGCAGACTCGCCAC	CTCTCTGGTGGGATGCAAC	2994	
3046	GGAGGTGTCCGTGGCCATTGCTTTGTGGGCGGCTCCCAA	AGTTGTTATCTTGGACGAGC	3105	
2995	GGAGGTGTCCGTGGCCATTGCTTTGTGGGCGGCTCCCAA	AGTTGTTATCTTGGACGAGC	3054	
3106	CTACGCTCGGCTGGATCTGTTCCCGCCGCGGTATTTGG	AGAGCTGTCTCAAAATACC	3165	
3055	CTACGCTCGGCTGGATCTGTTCCCGCCGCGGTATTTGG	AGAGCTGTCTCAAAATACC	3114	
3166	GAGAGGTGTGACGTGATCTCTCCACCCACCACTGGAT	TGAGCAGAGCTGTGGGAG	3225	
3115	GAGAGGTGTGACGTGATCTCTCCACCCACCACTGGAT	TGAGCAGAGCTGTGGGAG	3174	
3226	ACCGTGTGCTGTGTGGCAGGTGGCCGCTGTGTGTGTGT	GTGGTCCCGCACTTCTCTGC	3285	
3175	ACCGTGTGCTGTGTGGCAGGTGGCCGCTGTGTGTGTGT	GTGGTCCCGCACTTCTCTGC	3234	
3286	GCGCTCACCTGGGCTCCGCTACTACCTGAGCTGGTGA	AGCGCGGCTGTGCCCTGACCA	3345	
3235	GCGCTCACCTGGGCTCCGCTACTACTCTGAGCTGGTGA	AGCGCGGCTGTGCCCTGACCA	3294	

QY 3346 CCAATGAGAGGCTGACACTGACATGAGAGGCGAGTGTGGACACACAGGCGAGGAAAGAAGA 3405
DB |||||
QY 3295 CCAATGAGAGGCTGACACTGACATGAGAGGCGAGTGTGGACACACAGGCGAGGAAAGAAGA 3354
DB |||||
QY 3406 ATGGCAGCCAGGCGAGCAGAGTCCGCACTCCCTCAGCTGCTGCCCTGTGTCAGCACTGGG 3465
DB |||||
QY 3355 ATGGCAGCCAGGCGAGCAGAGTCCGCACTCCCTCAGCTGCTGCCCTGTGTCAGCACTGGG 3414
DB |||||
QY 3466 TGCCCGGGGCGACGGCTGTGTGGAGGAGCTGCCACACAGAGCTGTGTGCTGTGCCCTACA 3525
DB |||||
QY 3415 TGCCCGGGGCGACGGCTGTGTGGAGGAGCTGCCACACAGAGCTGTGTGCTGTGCCCTACA 3474
DB |||||
QY 3526 CGGTGCCCCATGACCGGCACTTTCGCCACACACTCTTCCGAGAGCTAGACACGGGCTGGCGG 3585
DB |||||
QY 3475 CGGTGCCCCATGACCGGCACTTTCGCCACACACTCTTCCGAGAGCTAGACACGGGCTGGCGG 3534
DB |||||
QY 3586 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGAGCTCGAGGAGATCTTCCCTGAAGG 3645
DB |||||
QY 3535 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAGGAGATCTTCCCTGAAGG 3594
DB |||||
QY 3646 TGGTGGAGGAGTGTCTCGGACACAGATATGGAGGATGGCAGCTGCGGGCAGACCTTAT 3705
DB |||||
QY 3595 TGGTGGAGGAGTGTCTCGGAGACAGATATGGAGGATGGCAGCTGCGGGCAGACCTTAT 3654
DB |||||
QY 3706 GCACAGGCATTGTCTGGCTAGACGTAACTCTCGGCTCAAGATGCCGCCACAGAGACAG 3765
DB |||||
QY 3655 GCACAGGCATTGTCTGGCTAGACGTAACTCTCGGCTCAAGATGCCGCCACAGAGACAG 3714
DB |||||
QY 3766 CGCTGGAGAACGGGGAACAGCTGGGTGAGCCCCAGACACTGACAGGGCTCTGGGCCAG 3825
DB |||||
QY 3715 CGCTGGAGAACGGGGAACAGCTGGGTGAGCCCCAGAGACTGACAGGGCTCTGGGCCAG 3774
DB |||||
QY 3826 ACGCGCTGGCGGGGTACAGGCTGGGCACTGACCCCGCAGCAGCTCCAGAGCCCTGCTTC 3885
DB |||||
QY 3775 ACGCGCTGGCGGGGTACAGGCTGGGCACTGACCCCGCAGCAGCTCCAGAGCCCTGCTTC 3834
DB |||||
QY 3886 TCAAGCGCTTTCGTCTGCCCGCGCAGCCCGCGGCTGTTCGCCCAGATCGTGTGC 3945
DB |||||
QY 3835 TCAAGCGCTTTCGTCTGCCCGCGCAGCCCGCGGCTGTTCGCCCAGATCGTGTGC 3894
DB |||||
QY 3946 CTGCGCTCTTTGTGGGCTGGCGCTCGTGTTCAGGCTCATGTGCTCTTTCGGGCACT 4005
DB |||||
QY 3895 CTGCGCTCTTTGTGGGCTGGCGCTCGTGTTCAGGCTCATGTGCTCTTTCGGGCACT 3954
DB |||||
QY 4006 ACCCGGCTCTGGGCTCAGTCCCACTGATGACGGTGTCTCAGGTGTCTTTCAGTGAGG 4065
DB |||||
QY 3955 ACCCGGCTCTGGGCTCAGTCCCACTGATGACGGTGTCTCAGGTGTCTTTCAGTGAGG 4014
DB |||||
QY 4066 ACGCCCCAGGGGACCCCTGGACGTGCCCGGCTCTCGAGGCGCTGTGCGAGGAGCGGAC 4125
DB |||||
QY 4015 ACGCCCCAGGGGACCCCTGGACGTGCCCGGCTCTCGAGGCGCTGTGCGAGGAGCGGAC 4074
DB |||||
QY 4126 TGGAGGAGCCCCAGTGACAGATAGCTCCCACTGCTTCAGGCTTCGGGACGAGAGTTCCTGCTG 4185
DB |||||
QY 4075 TGGAGGAGCCCCAGTGACAGATAGCTCCCACTGCTTCAGGCTTCAGGCTTCCTTTCAGTGAGG 4134
DB |||||
QY 4186 AAGTGGCCCAAGGCTTTGGGCCAGTGGCACTGAGCCACAGAGTCTCCATCCCGCAGCTGCC 4245
DB |||||
QY 4135 AAGTGGCCCAAGGCTTTGGGCCAGTGGCACTGAGCCCAAGATCTCCATCCCGCAGCTGCC 4194
DB |||||
QY 4246 AGTGTAGCAGCCCCGGTGGCGGCGCTGTGCCGACTGCGCGGCTGCAGCTGGTGGTC 4305
DB |||||
QY 4306 CCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGGTTCAGAACTTGACAGGCGGGA 4365
DB |||||
QY 4255 CCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGGTTCAGAACTTGACAGGCGGGA 4314
DB |||||
QY 4366 ACCTGTCTGACTTCTCTGTCTAAGACCTTACCGCGCTGGTGGCCAGGCGCTGAGACTA 4425
DB |||||
QY 4315 ACCTGTCTGACTTCTCTGTCTAAGACCTTACCGCGCTGGTGGCCAGGCGCTGAGACTA 4374
DB |||||

QY 4426 AGAAGTGGGTGAATGAGGTACGGATACGGAGGCTTCTCCCTGCGGGGGCCGAGACCCAGGCC 4485
DB |||||
QY 4375 AGAAGTGGGTGAATGAGGTACGGATACGGAGGCTTCTCCCTGCGGGGGCCGAGACCCAGGCC 4434
DB |||||
QY 4486 TGCCCTCGGGCCAGAGAGTTGGGCGGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCC 4545
DB |||||
QY 4435 TGCCCTCGGGCCAGAGAGTTGGGCGGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCC 4494
DB |||||
QY 4546 TGCCCTCGGGCGGCGCTCGACCGTGTCTCTGAAAAACCTCACAGCCTGGGCTCACAGCCTGG 4605
DB |||||
QY 4495 TGCCCTCGGGCGGCGCTCGACCGTGTCTCTGAAAAACCTCACAGCCTGGGCTCACAGCCTGG 4554
DB |||||
QY 4606 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTGGCT 4665
DB |||||
QY 4555 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTGGCT 4614
DB |||||
QY 4666 TTGTCAACCGAGCAGCAACGCAATCTCCGTGTCTCACTGCCCCCAGGCCCCGCCGCC 4725
DB |||||
QY 4615 TTGTCAACCGAGCAGCAACGCAATCTCCGTGTCTCACTGCCCCCAGGCCCCGCCGCC 4674
DB |||||
QY 4726 ACGCCCCAGCATCACACACTCAACACCCCTTTGAACCTCAACAAAGGAGCAGCTGTCTG 4785
DB |||||
QY 4675 ACGCCCCAGCATCACACACTCAACACCCCTTTGAACCTCAACAAAGGAGCAGCTGTCTG 4734
DB |||||
QY 4786 AGGCTGCACTGATGGCCTCTCGGTGAGAGTCTCTGCTTCCATCTGTGTGTCTTTGGCA 4845
DB |||||
QY 4735 AGGCTGCACTGATGGCCTCTCGGTGAGAGTCTCTGCTTCCATCTGTGTGTCTTTGGCA 4794
DB |||||
QY 4846 TGTCTTTTGTCCCGGCGCAGCTTCACTCTTGTCTCTATTGAGGAGGAGTCAACCGAGCA 4905
DB |||||
QY 4795 TGTCTTTTGTCCCGGCGCAGCTTCACTCTTGTCTCTATTGAGGAGGAGTCAACCGAGCA 4854
DB |||||
QY 4906 AGCACTGTGAGCTCATGGGGGCGCTGTCCCGCCACCTCTACTGTGCTTGGCAACTTCTCT 4965
DB |||||
QY 4855 AGCACTGTGAGCTCATGGGGGCGCTGTCCCGCCACCTCTACTGTGCTTGGCAACTTCTCT 4914
DB |||||
QY 4966 GGGACATGTGTAACTATTTGTTGCCAGCATGCTGTGTGTCTCATCTTTCTGCGCTTCC 5025
DB |||||
QY 4915 GGGACATGTGTAACTATTTGTTGCCAGCATGCTGTGTGTCTCATCTTTCTGCGCTTCC 4974
DB |||||
QY 5026 AGCAGAGGGATATGTGGCCCCCGCCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5085
DB |||||
QY 4975 AGCAGAGGGATATGTGGCCCCCGCCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 5034
DB |||||
QY 5086 ATGCTGTGTGCTCATCACACCGCTCATGTACCCAGGCTCTCTCTCTCTCTCTCTCTCTCT 5145
DB |||||
QY 5035 ATGCTGTGTGCTCATCACACCGCTCATGTACCCAGGCTCTCTCTCTCTCTCTCTCTCTCT 5094
DB |||||
QY 5146 CAGCTATGTGTGCTCACCTGCAATAAACCTCTTTTATTGGCATCAATGGAAAGCATGGCCA 5205
DB |||||
QY 5095 CAGCTATGTGTGCTCACCTGCAATAAACCTCTTTTATTGGCATCAATGGAAAGCATGGCCA 5154
DB |||||
QY 5206 CTTTGTGTGTGAGCTCTTCTCTGATCAGAAGCTGCGAGGAGGTGAGCGGATCTTGAAC 5265
DB |||||
QY 5155 CTTTGTGTGTGAGCTCTTCTCTGATCAGAAGCTGCGAGGAGGTGAGCGGATCTTGAAC 5214
DB |||||
QY 5266 AGGTCTTCTTATCTTCTCCCACTTCTGTCTGGGCGCGGGGCTCATATTGACATGGTCGGA 5325
DB |||||
QY 5215 AGGTCTTCTTATCTTCTCCCACTTCTGTCTGGGCGCGGGGCTCATATTGACATGGTCGGA 5274
DB |||||
QY 5326 ACCAGGCCATGGCTGTATGCTCTTTGAGCGCTTTGGAGACAGGACAGTTCCAGTCACCCCCTGC 5385
DB |||||
QY 5275 ACCAGGCCATGGCTGTATGCTCTTTGAGCGCTTTGGAGACAGGACAGTTCCAGTCACCCCCTGC 5334
DB |||||
QY 5386 GCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGTATACAGGGGCGCTCTTCTCTTC 5445
DB |||||
QY 5335 GCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGTATACAGGGGCGCTCTTCTCTTC 5394
DB |||||
QY 5446 TCTTCACTACTGCTCAGCACCGAAGCCAACTCTCTGCGCCACAGCCAGGCTGAGGTCTC 5505
DB |||||
QY 5395 TCTTCACTACTGCTCAGCACCGAAGCCAACTCTCTGCGCCACAGCCAGGCTGAGGTCTC 5454
DB |||||
QY 5506 TGCCACTCTCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAACGGGAGCGGCTGCTCAAG 5565
DB |||||

5455	Db		TGCCACTCTCTGGGAGAGGAGCAGAGATATAGCCCTGAAACGGGAGCGGTGGTCCAAG	5514
5566	Qy		GAGCCACCCAGGGGGATGTGTGGTGTGAGGAACATTGACCAAGGTATATCCGTGGGCAGA	5625
5515	Db		GAGCCACCCAGGGGGATGTGTGGTGTGAGGAACATTGACCAAGGTATATCCGTGGGCAGA	5574
5626	Qy		GGATGCCAGCTGTTGACCGCTTGTGCTCGGGGATTCGCCCTGTGTGAGTGTGTTTGGGCTGC	5685
5575	Db		GGATGCCAGCTGTTGACCGCTTGTGCTCGGGGATTCGCCCTGTGTGAGTGTGTTTGGGCTGC	5634
5686	Qy		TGGGTGTGAATGAGCAGGGAAGACGTCACGTTTGCATTGGTGAAGGGGACACATTGG	5745
5635	Db		TGGGTGTGAATGAGCAGGGAAGACGTCACGTTTGCATTGGTGAAGGGGACACATTGG	5694
5746	Qy		CCAGCAGGGGCGAGGCTGTCTGGCAGGCCACACAGCTGGGCCCGGGAACCCAGTGTCTGCGC	5805
5695	Db		CCAGCAGGGGCGAGGCTGTCTGGCAGGCCACACAGCTGGGCCCGGGAACCCAGTGTCTGCGC	5754
5806	Qy		ACTCAGCATGGGATATCGCCCTCAATCCGATCGCATCTTTTGAGCTGTCTGACGGGCGCG	5865
5755	Db		ACTCAGCATGGGATATCGCCCTCAATCCGATCGCATCTTTTGAGCTGTCTGACGGGCGCG	5814
5866	Qy		AGCACTTGGAGCTCTTTCGCGCGCTCGCGGCTGTCCGGAGGCCCAAGTTGCCACAGACCG	5925
5815	Db		AGCACTTGGAGCTCTTTCGCGCGCTCGCGGCTGTCCGGAGGCCCAAGTTGCCACAGACCG	5874
5926	Qy		CTGGCTCAGGCTTGGCGCGTCTGGGACTCTCATGTGTACGACAGACCGGCTGTGAGGACCT	5985
5875	Db		CTGGCTCAGGCTTGGCGCGTCTGGGACTCTCATGTGTACGACAGACCGGCTGTGAGGACCT	5934
5986	Qy		ACAGCGAGGGAAACAAACGAACTGCGCAGCGCCCTGGCGCTGTGTTGGGGACCCAGACCG	6045
5935	Db		ACAGCGAGGGAAACAAACGAACTGCGCAGCGCCCTGGCGCTGTGTTGGGGACCCAGACCG	5994
6046	Qy		TGTTGTTTCTGGACGACCGACCAACAGGCACTGGAACCCACGCGCGCGCGCTTCTTTGGA	6105
5995	Db		TGTTGTTTCTGGACGACCGACCAACAGGCACTGGAACCCACGCGCGCGCGCTTCTTTGGA	6054
6106	Qy		ACAGCTTTTGGCGGTGGTCGGAGGGCGGTTTCAGTGATGCTCACTCCCATAGCATGG	6165
6055	Db		ACAGCTTTTGGCGGTGGTCGGAGGGCGGTTTCAGTGATGCTCACTCCCATAGCATGG	6114
6166	Qy		AGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTTGAATGGGCGGTTCCGCTGCC	6225
6115	Db		AGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTTGAATGGGCGGTTCCGCTGCC	6174
6226	Qy		TGGGACGCCCGCAACATCTCAAGGGCAGATTCCGGCGGGTCAACACTGACCTCGGG	6285
6175	Db		TGGGACGCCCGCAACATCTCAAGGGCAGATTCCGGCGGGTCAACACTGACCTCGGG	6234
6286	Qy		TGCCCGCCGCAAGGTCCACGCGGACGCGCTTGTGTGGCGGCGGATTCCTTGGGTCCG	6345
6235	Db		TGCCCGCCGCAAGGTCCACGCGGACGCGCTTGTGTGGCGGCGGATTCCTTGGGTCCG	6294
6346	Qy		AGCTGCGCAGGGACATGAGAGCGCTCGCTTCAGAGTCCGCGCGGAGGCGCTGCG	6405
6295	Db		AGCTGCGCAGGGACATGAGAGGTTCGCTCGCTTCAGAGTTCGCGCGGAGGCGCTGCG	6354
6406	Qy		CCCTGGCGCGCTCTTTTGGAGAGCTGGCGGTGCAACGGCGCAGAGCAACGGCTGGAGACT	6465
6355	Db		CCCTGGCGCGCTCTTTTGGAGAGCTGGCGGTGCAACGGCGCAGAGCAACGGCTGGAGACT	6414
6466	Qy		TTTCCGTGAGCCAGACGATGCTGGAGAGTATTTCTTGTACTTCTCAAGGACAGGGGA	6525
6415	Db		TTTCCGTGAGCCAGACGATGCTGGAGAGTATTTCTTGTACTTCTCAAGGACAGGGGA	6474
6526	Qy		AGGACAGGACACCGAAGACAGAAAGAGCAGGAGTGGGAGTGGACCCCGCCAGGCC	6585
6475	Db		AGGACAGGACACCGAAGACAGAAAGAGCAGGAGTGGGAGTGGACCCCGCCAGGCC	6534
6586	Qy		TGCAGCACCCCAACCGCGTCAGCCAGTTCCTCGATGACCTTAGCACTTGCACGAGCTGTGC	6645

DEC 11 1964

RESULT 4
ADD37439

ADD37428
ID ADD37428 standard: cDNA: 6768 BP.

AC ADD37428:

DT 15--JAN-2004 (first entry)

XX
DE Human transporter MTP-1 cDNA #1.

Human; ss; gene; transporter; cytostatic; anorectic; antidiabetic;
KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;
KW obesity; epilepsy; diabetes.

Homogamietisch.

XX PN US2003143675-A1.

31-июл.-2003.

22-MAY-2002: 2002US-00154419.

PR 12-MAY-2000: 2000US-0204211P.

23-JUN-2000; 2000US-0213370E:
PR 31-JUL-2000: 2000US-0221769P:
PR

PR 25-SEP-2000: 2000US-0235107P.

PR 14-NOV-2000; 2000US-0248364P.

PR 15-DEC-2000; 2000US-0256240P.

21-DEC-2000: 2000US-0258028P.

PR 14-MAY-2001; 2001US-00858194.

PR 31-JUL-2001: 2001US-00919781.

PR 25-SEP-2001; 2001US-00964295.

PR 14-NOV-2001: 2001US-00002769:

22-JAN-2002; 2002US-00055025.

PA (MILL-) MILLENNIUM PHARM INC.

PI Curtis RAJ, Glucksmann MA,

DR WPI; 2003-851783/79.

XX
XX

PT treating PGC-1 associated dis

1999

XXXXXX

.....

(MILL-) MILLENNIUM PHARM INC.

Curtis RAJ, Glucksmann MA, Meyers RE;

WPI; 2003-851783/79.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy or diabetes.

Claim 1: SEQ ID NO 1; 663pp; English.

The invention relates to an isolated nucleic acid comprising a cDNA

encoding a human transporter protein, or its complement, a sequence that is 60 % identical to the cDNA, a fragment comprising at least 30 nucleotides of the cDNA, or a sequence encoding a fragment of the polypeptide comprising at least 10 contiguous amino acid residues of the cDNA. Also included are a vector comprising the novel nucleic acid molecule, producing the polypeptide, the isolated transporter polypeptide, an isolated antibody that specifically binds to the polypeptide, detecting the presence of the polypeptide or nucleic acid in a sample, a kit, identifying a compound that binds to, or that modulates the activity of, the polypeptide, and modulating the activity of the polypeptide. The nucleic acid is useful for preparing a composition for treating PCG-1 (not defined) associated disorders e.g. liver tumours, obesity, epilepsy or diabetes. The present sequence encodes a novel human transporter protein.

SQ Sequence 6768 BP; 1143 A; 2190 C; 2170 G; 1265 T; 0 U; 0 Other;

Query Match 98.7%; Score 6717.8; DB 10; Length 6768;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 6741; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

QY 46 CGCAAGAGCTGCGAGCCCTGGAAGAGCTTCCAGGAACCTGCCGTGTGGGATAAAGG 105

Db
1 CGGCAAGAGCTGCGGAGCCCTGGAAGAGCTTCCAGGAACCTGCGCTGTGGATAAAGG 60

106 AATGAGGTTCA GAAAGGGGCAGGGAGTTGCCCGCAGCGGACCGGACCGTCTTCAGCCCCA 165

61

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~
_

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

GG
12T CCGTAAGTCCCTGACCCCTCCTGTCCCCCTGCTCCCTCAGTCCTCACCAATGGCCCTTCCTGGACAC 180

QY
226 AGCTGATGCTGCTGCTCTGGAAAGAAATTTCATGTAATCGCCGGAGACAGCCGGTCCAGCTCC 285

Db 181 AGCTGATGCTGCTCTGGAAGAATTTTCATGTATCGCCGGAGACAGCC-----GCTCC 234

286 TGGTCGAATTGCTGTGGCCCTCTCTTCCCTCTTCTTCATCCTGGTGGCTGTTGCCACTCCC 345

D_b 235 TGGTCGAATTGCTGTGGCCCTCTCTTCCCTCTTCTTCATCCTGGTGGCTGTTCGCCACTCCC 294

QY 346 ACCCGCCCCCTGGAGCACCATGAATGCCACTTCCCAAACAAGCCACTGCCATCGGCGGGCA 405

Db
295 ACCCGCCCTGGAGCACCATGAATGCCACTTCCCAACAAGCCACTGCCATCGCGGGCA 354

OV 406 CCGTGCCCTGGCTCCAGGGTCTCATCTGTAAATGTGAACAACACTGCTTTCGGCAGGTGA 465

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 328 TAGCCGATGCCCCGACATGTCATGGAGGGGCTCAGTGTCCACAGGACGCTGGCTGGCCTAG 585

DB 4/5 TAGCCGATGCCCGCACCTGTGCTGGGAGGGGCCAGTGCCCA CAGGACGCTGGCTGGCCTAG 534

Qy 586 GGAAGCTGATCGCCACGCTGAGGGCTGCACGCAGCAGGCCCTCAACCAACCAAGC 645

Db 535 GGAAGCTGATCGCCACGCTGAGGGCTGCACGCAGCACGGCCAGCCTCAACCAACCAGC 594

QY 646 AGTCTCCACTGGAACCAACCCATGCTGGATGTCGCGGAGCTGCTGACGTCACCTGCTGCGCA 705

Db 595 AGTCTCCACTGGAAACCAACCCATGCTGGATGTCGCGGAGCTGCTGACGTCACTGTCTGCGCA 654

QV 706 CGGAATCCCTGGGGTTGGCACTGGGCCAAGCCCAGGAGCCCTTGCA CAGCTTGTTGGAGG 765

Db
655 CGGAATCCCTGGGGTTGGCACTGGGCCAAGCCAGGAGCCCTTGCACAGCTTGTCGAGG 714

Db	1855	TGAAGGCGGTGGTGGGAGAGAGAGCGGGCTCGGGACACCATGCGCGCCATGGGC	1914
Qy	1966	TCAGCCGCGGGTCTCTGGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTGCTCA	2025
Db	1915	TCAGCCGCGGGTCTCTGGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTGCTCA	1974
Qy	2026	GGCGCGGCTGCTGGTTCCTGGTCTCAAGCTGGGGACATCCCTCCCTACAGCACCCGG	2085
Db	1975	GGCGCGGCTGCTGGTTCCTGGTCTCAAGCTGGGGACATCCCTCCCTACAGCACCCGG	2034
Qy	2086	GGCTGGTCTTCCTGTTCTTGGCAGCCTTCGCGGTGGCCACGCTGACCCAGAGCTTCCTGC	2145
Db	2035	GGCTGGTCTTCCTGTTCTTGGCAGCCTTCGCGGTGGCCACGCTGACCCAGAGCTTCCTGC	2094
Qy	2146	TCAGCGCTTCCTTCCTCCGCGCCAACTCGGTGGGCTCGGGGCTCGGGGCTCGGCTCTCT	2205
Db	2095	TCAGCGCTTCCTTCCTCCGCGCCAACTCGGTGGGCTCGGGGCTCGGGGCTCGGCTCTCT	2154
Qy	2206	CCCTCTACCTGCCCTACGTGCTGTGTGGCTTGGCGGGACCGGCTGCGCGGGTGGCC	2265
Db	2155	CCCTCTACCTGCCCTACGTGCTGTGTGGCTTGGCGGGACCGGCTGCGCGGGTGGCC	2214
Qy	2266	GCCTGGCGGAGGCTCTGTTCGCGCTGGCCCTTCGGCTTCGGCTGGAGAGCTTCGGCTC	2325
Db	2215	GCCTGGCGGAGGCTCTGTTCGCGCTGGCCCTTCGGCTTCGGCTGGAGAGCTTCGGCTC	2274
Qy	2326	TGCTGGAGAGCAGGGCGAGGCGCGCAGTGGGACAACTGGGGACCCCGGCTTCAGGCG	2385
Db	2275	TGCTGGAGAGCAGGGCGAGGCGCGCAGTGGGACAACTGGGGACCCCGGCTTCAGGCG	2334
Qy	2386	ACGTCTTCAGCCTGGCCAGGCTCTCGGCTTCGTGCTGGAGCGGGCGCTCTACGGCC	2445
Db	2335	ACGTCTTCAGCCTGGCCAGGCTCTCGGCTTCGTGCTGGAGCGGGCGCTCTACGGCC	2394
Qy	2446	TCGCCACCTGGTACTTGGAAGCTGTGTGCCAGGCCAGTAGTACGGGATCCCTGAAACCATGGA	2505
Db	2395	TCGCCACCTGGTACTTGGAAGCTGTGTGCCAGGCCAGTAGTACGGGATCCCTGAAACCATGGA	2454
Qy	2506	ATTTTCCTTTTCGGAGAGCTACTGGTGGGACCTTCGGCCCCCAGAGTCCAGCCCTT	2565
Db	2455	ATTTTCCTTTTCGGAGAGCTACTGGTGGGACCTTCGGCCCCCAGAGTCCAGCCCTT	2514
Qy	2566	GGCCACCCGCTGGACCCAAAGTGTGTAGAGAGGACCGCCGGCTCAGTCCCTG	2625
Db	2515	GGCCACCCGCTGGACCCAAAGTGTGTAGAGAGGACCGCCGGCTCAGTCCCTG	2574
Qy	2626	GCCTATCCGTTTCGAGCCTGGAGAGCGCTTTCCTGGAAGCCCGCAGCCAGCCCTCGGG	2685
Db	2575	GCCTATCCGTTTCGAGCCTGGAGAGCGCTTTCCTGGAAGCCCGCAGCCAGCCCTCGGG	2634
Qy	2686	GGCTACGCTGGACTTCTACAGGGCCACATCACCGCTTCCTGGGGCCAAAGGGGGCG	2745
Db	2635	GGCTACGCTGGACTTCTACAGGGCCACATCACCGCTTCCTGGGGCCAAAGGGGGCG	2694
Qy	2746	GCAGACACACCTCTGCATCTTAGTGGCCCTTCCTCCACCCAGTGGGCTCTGCCT	2805
Db	2695	GCAGACACACCTCTGCATCTTAGTGGCCCTTCCTCCACCCAGTGGGCTCTGCCT	2754
Qy	2806	TCATCTCTGGGCCACAGCCTCGGCTCCAGATGGCGGCCATCCCGCCCACTCGGGGCTCT	2865
Db	2755	TCATCTCTGGGCCACAGCCTCGGCTCCAGATGGCGGCCATCCCGCCCACTCGGGGCTCT	2814
Qy	2866	GTCTCTAGTCAACAGTCTGTTGACATGCTGACCGTGGAGCAGAGCTCTGTTCTATG	2925
Db	2815	GTCTCTAGTCAACAGTCTGTTGACATGCTGACCGTGGAGCAGAGCTCTGTTCTATG	2874
Qy	2926	GGCGGCTGAAGGCTCTCAGTGGCGCTGAGTGGGCCCCCGAGCAGGACCGTCTCGAGG	2985
Db	2875	GGCGGCTGAAGGCTCTCAGTGGCGCTGAGTGGGCCCCCGAGCAGGACCGTCTCGAGG	2934
Qy	2986	ATGTGGGCTGGTCTCAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATGCAAC	3045

Db	2935	ATGTGGGCTGGTCTCAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATGCAAC	2994
Qy	3046	GGAGCTGTCCGTGGGCCATTTGCTTTGGCGGCTCCCAAGTTGTATTCCTGGAGGAC	3105
Db	2995	GGAGCTGTCCGTGGGCCATTTGCTTTGGCGGCTCCCAAGTTGTATTCCTGGAGGAC	3054
Qy	3106	CTACGGCTGGCGTGGATCTGCTTCCCGCGCGGTTATTTGGGAGCTGCTGCTCAAAATACC	3165
Db	3055	CTACGGCTGGCGTGGATCTGCTTCCCGCGCGGTTATTTGGGAGCTGCTGCTCAAAATACC	3114
Qy	3166	GAGAAGTCCGACGCTGATCTCTCCACCCACCACTGATGAGGAGAGCTGCTGGAG	3225
Db	3115	GAGAAGTCCGACGCTGATCTCTCCACCCACCACTGATGAGGAGAGCTGCTGGAG	3174
Qy	3226	ACCGTGTGGCTGTGGTGGCAGGTGGCGGCTGTGCTGCTGTGGCTCCCACTTCTCTGC	3285
Db	3175	ACCGTGTGGCGGTGGTGGCAGGTGGCGGCTGTGCTGCTGTGGCTCCCACTTCTCTGC	3234
Qy	3286	GCCTCTCACCTGGGCTCCGGCTACTACTGACGCTGATGAAGGCCCGCTTGCCTTACCA	3345
Db	3235	GCCTCTCACCTGGGCTCCGGCTACTACTGACGCTGATGAAGGCCCGCTTGCCTTACCA	3294
Qy	3346	CCAATCAGAAAGCTGACATGACATGAGGGGCACTGTGACACACGAGGAGGAGGAGGAG	3405
Db	3295	CCAATCAGAAAGCTGACATGACATGAGGGGCACTGTGACACACGAGGAGGAGGAGGAG	3354
Qy	3406	ATGGCAGCCAGGGCAGAGGCTCGGCACTCTCAGCTGTGGCCCTGTGTACAGCACTGGG	3465
Db	3355	ATGGCAGCCAGGGCAGAGGCTCGGCACTCTCAGCTGTGGCCCTGTGTACAGCACTGGG	3414
Qy	3466	TGCCCCGGGGCAGCGGCTGGTGGAGGAGCTGCACACAGCTGTGTGTGTGTGTGTGTGTGT	3525
Db	3415	TGCCCCGGGGCAGCGGCTGGTGGAGGAGCTGCACACAGCTGTGTGTGTGTGTGTGTGT	3474
Qy	3526	CGGGTCCCATGACGCGAGCTTTCGCCACACTCTTCCGAGAGCTAGACACGCGCTGGCGG	3585
Db	3475	CGGGTCCCATGACGCGAGCTTTCGCCACACTCTTCCGAGAGCTAGACACGCGCTGGCGG	3534
Qy	3586	AGCTAGGCTCACTGGCTAGGGATCTCCGACACAGCTTCGAGGAGATCTTCTGTGAGG	3645
Db	3535	AGCTAGGCTCACTGGCTAGGGATCTCCGACACAGCTTCGAGGAGATCTTCTGTGAGG	3594
Qy	3646	TGCTGAGGAGTGTGTCTGCGGACACAGATATGAGGATGGCAGCTCGGGCAGCACTAT	3705
Db	3595	TGCTGAGGAGTGTGTCTGCGGACACAGATATGAGGATGGCAGCTCGGGCAGCACTAT	3654
Qy	3706	GCAAGGCAATTTGCTGCGCTTAGAGTAACCTTCGCGCTCAAGATGCGCCACAGAGACAG	3765
Db	3655	GCAAGGCAATTTGCTGCGCTTAGAGTAACCTTCGCGCTCAAGATGCGCCACAGAGACAG	3714
Qy	3766	CGCTGAGAACGGGGAACGAGCTGGGTACGCCAGAGCTGACAGGCTCTGGGCGCAG	3825
Db	3715	CGCTGAGAACGGGGAACGAGCTGGGTACGCCAGAGCTGACAGGCTCTGGGCGCAG	3774
Qy	3826	ACGCCCTGGGCGGGGTACAGGGCTGGGCTGACCCCGCAGAGCTCCAGGCGCTGTTC	3885
Db	3775	ACGCCCTGGGCGGGGTACAGGGCTGGGCTGACCCCGCAGAGCTCCAGGCGCTGTTC	3834
Qy	3886	TCAAGCGCTTTCTGCTTTCGCGCGCGCAGCGCGCGCTGTTCGCGCGAGATCGTGTGC	3945
Db	3835	TCAAGCGCTTTCTGCTTTCGCGCGCGCAGCGCGCGCTGTTCGCGCGAGATCGTGTGC	3894
Qy	3946	CTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCTCATCGTGTCTCTTTCGGGCACT	4005
Db	3895	CTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCTCATCGTGTCTCTTTCGGGCACT	3954
Qy	4006	ACCGGCTCTGCGGCTCAGTCCCACTGATGAGGCTCAGGTGTCTTCTTTCAGTGAGG	4065
Db	3955	ACCGGCTCTGCGGCTCAGTCCCACTGATGAGGCTCAGGTGTCTTCTTTCAGTGAGG	4014
Qy	4066	ACGCCCAAGGGGACCTCGACGCTGCGCGCTGTCTGAGGCGCTGTCTGAGGAGGAGGAC	4125
Db	4015	ACGCCCAAGGGGACCTCGACGCTGCGCGCTGTCTGAGGCGCTGTCTGAGGAGGAGGAC	4074

QY	4126	TGGAGAGGCCCCCAGTGCAGCATAGTCTCCACAGGTTCTCGCACCGAGAAGTTCTGTCTG	4185
Db	4075	TGGAGAGGCCCCCAGTGCAGCATAGTCTCCACAGGTTCTCGCACCGAGAAGTTCTGTCTG	4134
QY	4186	AAAGTGGCCAAAGTCTTTGGCCAGTGGCACTGAACCCAGAGTCTCCATCCCGAGCTGCC	4245
Db	4135	AAAGTGGCCAAAGTCTTTGGCCAGTGGCACTGAACCCAGAGTCTCCATCCCGAGCTGCC	4194
QY	4246	AGTGTAGCCAGCCCGGTGCCCGGCGCTGTCTCCCGACTGCCCGGCTGCAGCTGTGTGTC	4305
Db	4195	AGTGTAGCCAGCCCGGTGCCCGGCGCTGTCTCCCGACTGCCCGGCTGCAGCTGTGTGTC	4254
QY	4306	CCCCCTCCGCCAGCAGAGTGAACCGGCTCTGGGGAAGTGGTTGAGAACCTGACAGCCCGGA	4365
Db	4255	CCCCCTCCGCCAGCAGAGTGAACCGGCTCTGGGGAAGTGGTTGAGAACCTGACAGCCCGGA	4314
QY	4366	ACCTGTCTGACTTCTTGCTCAAGACTACCCCGGCTGTGGTGGCCAGGGCTTGAAGACTA	4425
Db	4315	ACCTGTCTGACTTCTTGCTCAAGACTACCCCGGCTGTGGTGGCCAGGGCTTGAAGACTA	4374
QY	4426	AGAACTGGGTGAATCAGGTACAGGTACGAGGCTTCTCGCTGGGGGGCGAGACCCAGGCC	4485
Db	4375	AGAACTGGGTGAATCAGGTACAGGTACGAGGCTTCTCGCTGGGGGGCGAGACCCAGGCC	4434
QY	4486	TGCCCTCGGGGCCAAGAGTTGGGCCGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGTCCCC	4545
Db	4435	TGCCCTCGGGGCCAAGAGTTGGGCCGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGTCCCC	4494
QY	4546	TGCCCTGGGGGGCCCTCGACCGTGTCTTGAAAACTCAGAGCTGGGGCTCAGAGCCTGG	4605
Db	4495	TGCCCTGGGGGGCCCTCGACCGTGTCTTGAAAACTCAGAGCTGGGGCTCAGAGCCTGG	4554
QY	4606	ATGCTCAGCAGACTCTCAAGTCTGTTTCAACAACAAAGGCTGGCACTCCATGTGGCCT	4665
Db	4555	ATGCTCAGCAGACTCTCAAGTCTGTTTCAACAACAAAGGCTGGCACTCCATGTGGCCT	4614
QY	4666	TTGTCAACCGAGCCAGCAACCAATCTCCGTGTCTCAGCTGCCCGCCGCGCCGCGCC	4725
Db	4615	TTGTCAACCGAGCCAGCAACCAATCTCCGTGTCTCAGCTGCCCGCCGCGCCGCGCC	4674
QY	4726	AGGCCACAGCATACCACTCAACCAACCCCTTGAACTCAACCAAGAGAGCTGTGTG	4785
Db	4675	AGGCCACAGCATACCACTCAACCAACCCCTTGAACTCAACCAAGAGAGCTGTGTG	4734
QY	4786	AGGCTGCACTGATGGCCCTCTCGGTGGAGCTCCTCGTCTCCATCTGTGTGTCTTTGCCA	4845
Db	4735	AGGCTGCACTGATGGCCCTCTCGGTGGAGCTCCTCGTCTCCATCTGTGTGTCTTTGCCA	4794
QY	4846	TGTCCTTTGTCCCGGCCAGCTTCACTCTGTGTCTCATTTGAGAGCGAGTCAACCGAGCCA	4905
Db	4795	TGTCCTTTGTCCCGGCCAGCTTCACTCTGTGTCTCATTTGAGAGCGAGTCAACCGAGCCA	4854
QY	4906	AGCACTTCGAGCTCATAGGGGGCGCTGTCCCGCCCTCTACTTGGCTTGGCAACTTTCTCT	4965
Db	4855	AGCACTTCGAGCTCATAGGGGGCGCTGTCCCGCCCTCTACTTGGCTTGGCAACTTTCTCT	4914
QY	4966	GGGACATGTGTAACTACTTGTGCGCAGCATGATGGTGTCTCATCTTTCTGGCCTTCC	5025
Db	4915	GGGACATGTGTAACTACTTGTGCGCAGCATGATGGTGTCTCATCTTTCTGGCCTTCC	4974
QY	5026	AGCAGAGGGCATATGTGGCCCTCGCAACCTGCGCTCTCTCGCTGTGTGTACTACTGT	5085
Db	4975	AGCAGAGGGCATATGTGGCCCTCGCAACCTGCGCTCTCTCGCTGTGTGTACTACTGT	5034
QY	5086	ATGGCTGTGTCATCACACCGCTCATGTATACCGAGCCTCTCTTTCTTCGTCGCGCCAGCA	5145
Db	5035	ATGGCTGTGTCATCACACCGCTCATGTATACCGAGCCTCTCTTTCTTCGTCGCGCCAGCA	5094
QY	5146	CAGCCTATGTGGTGTCTACCTGTCATAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCA	5205
Db	5095	CAGCCTATGTGGTGTCTACCTGTCATAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCA	5154

QY	5206	CCTTTGTGCTTGTAGCTCTTCTGTGATCAGAAAGCTCAGAGAGTGTAGCCGGATCTTTGAAAC	5265
Db	5155	CCTTTGTGCTTGTAGCTCTTCTGTGATCAGAAAGCTCAGAGAGTGTAGCCGGATCTTTGAAAC	5214
QY	5266	AGGTCTTCTTATCTTCCCGCCACTTCTGTCTTTGGGCGGGGGCTCATTTGACATGTTGCCGA	5325
Db	5215	AGGTCTTCTTATCTTCCCGCCACTTCTGTCTTTGGGCGGGGGCTCATTTGACATGTTGCCGA	5274
QY	5326	ACCAAGCCATGGCTGATGCCCTTTGAGCGCTTTGGGAGACAGGCAGTTTCAGTCAACCCCTGC	5385
Db	5275	ACCAAGCCATGGCTGATGCCCTTTGAGCGCTTTGGGAGACAGGCAGTTTCAGTCAACCCCTGC	5334
QY	5386	GCTGGAGGTGGTCCGCAAGAACCTCTTTGGCCATGTTGATACAGGGGCCCTCTTCTCTTC	5445
Db	5335	GCTGGAGGTGGTCCGCAAGAACCTCTTTGGCCATGTTGATACAGGGGCCCTCTTCTCTTC	5394
QY	5446	TCCTTCACTACTGCTGAGCAGCACCGAAGCCAACTCTCTGCCACAGCCACAGGGTGTGCTC	5505
Db	5395	TCCTTCACTACTGCTGAGCAGCACCGAAGCCAACTCTCTGCCACAGCCACAGGGTGTGCTC	5454
QY	5506	TGCCACTCTCGGAGAGGAGCAGAGGATGTAGCCCGTGAACGGGAGCGGGTGTGCTCAAG	5565
Db	5455	TGCCACTCTCGGAGAGGAGCAGAGGATGTAGCCCGTGAACGGGAGCGGGTGTGCTCAAG	5514
QY	5566	GAGCCACCCAGGGGGATGTGTGTGTCTGAGGAACTTGAACCAAGGTATACCGTGGGCGAGA	5625
Db	5515	GAGCCACCCAGGGGGATGTGTGTGTCTGAGGAACTTGAACCAAGGTATACCGTGGGCGAGA	5574
QY	5626	GGATGCCAGCTTGTGAACCGCTTGTCTGGGGATTTCCCTCTGTGTGAGTGTTTTGGGCTGC	5685
Db	5575	GGATGCCAGCTTGTGAACCGCTTGTCTGGGGATTTCCCTCTGTGTGAGTGTTTTGGGCTGC	5634
QY	5686	TGGGTGTGAATGGAGCAGGGAAGACGCTCCAGTTTGCATGCTGACGGGGGACACATTTGG	5745
Db	5635	TGGGTGTGAATGGAGCAGGGAAGACGCTCCAGTTTGCATGCTGACGGGGGACACATTTGG	5694
QY	5746	CCAGCAGGGGCGAGGCTGTGTCTGCGAGGCCACAGCGTGGCCCGGGAAACCCAGTGTGCGC	5805
Db	5695	CCAGCAGGGGCGAGGCTGTGTCTGCGAGGCCACAGCGTGGCCCGGGAAACCCAGTGTGCGC	5754
QY	5806	ACCTCAGCATGGGATATCTGCCCTCAATCCGATGCGCATTTTGTGCTGTGACGGGCGCG	5865
Db	5755	ACCTCAGCATGGGATATCTGCCCTCAATCCGATGCGCATTTTGTGCTGTGACGGGCGCG	5814
QY	5866	AGCACTTGGAGCTGTTGCGCGCTGCGCGGTGTCGGGAGGCCAGGTTGCCACAGACCG	5925
Db	5815	AGCACTTGGAGCTGTTGCGCGCTGCGCGGTGTCGGGAGGCCAGGTTGCCACAGACCG	5874
QY	5926	CTGGCTCAGGCTTGGCGCTCTGGGACTCTCATGTGTAGCAGACCGGCTTGAGGCACCT	5985
Db	5875	CTGGCTCAGGCTTGGCGCTCTGGGACTCTCATGTGTAGCAGACCGGCTTGAGGCACCT	5934
QY	5986	ACAGCGAGGGAACAAACGCAAGTGGCGACCGCCCTTGGCGCTGTGTTGGGACCCAGCCG	6045
Db	5935	ACAGCGAGGGAACAAACGCAAGTGGCGACCGCCCTTGGCGCTGTGTTGGGACCCAGCCG	5994
QY	6046	TGGTGTCTTGGACGAGCCGACCAAGGATGAGACCCAGCGCGGGGGCTTCTTTTGA	6105
Db	5995	TGGTGTCTTGGACGAGCCGACCAAGGATGAGACCCAGCGCGGGGGCTTCTTTTGA	6054
QY	6106	ACAGCCTTTTGGCGTGTGGGAGGCGCTTTCAGTGTATGCTCACTCCCATAGCATGG	6165
Db	6055	ACAGCCTTTTGGCGTGTGGGAGGCGCTTTCAGTGTATGCTCACTCCCATAGCATGG	6114
QY	6166	AGGAGTGTGAAGCGCTCTGTCTCGCGCTTAGCCATCATGTGTAATGGCGGTTCCGCTGCC	6225
Db	6115	AGGAGTGTGAAGCGCTCTGTCTCGCGCTTAGCCATCATGTGTAATGGCGGTTCCGCTGCC	6174
QY	6226	TGGCAGCCCGCAACATCTCAAGGGGAGATTCGGGGGGGGTCAACACTGACCTTGGGG	6285
Db	6175	TGGCAGCCCGCAACATCTCAAGGGGAGATTCGGGGGGGGTCAACACTGACCTTGGGG	6234
QY	6286	TGCCCGCGGCAAGGTCCAGCCGCGGCGCTTCTGTGGCGCGGAGTTCCCTGGGTGG	6345

Db 6235 TGGCCGCGCAAGTCCACGCGGACGCGCTTCGTGGCGCGGAGTTCCTGGTGG 6294
Qy 6346 AGTGGCGGAGGACATGAGAGCGCGCTGCGCTTCAGCTGCGCGGAGAGCGCTGGG 6405
Db 6295 AGCTGGCGGAGGACATGAGAGTTCGCTGCGCTTCAGCTGCGCGGAGAGCGCTGGG 6354
Qy 6406 CCTGGCGCGCTTCCTGGAGAGCTGGCGGTGACCGCGGACAGACGCGGTGGAGACT 6465
Db 6355 CCTGGCGCGCTTCCTGGAGAGCTGGCGGTGACCGCGGACAGACGCGGTGGAGACT 6414
Qy 6466 TTTCCGCTGAGCAGACATGCTGGAGAGGTATTTCTGTACTTCTCAAGGACACAGGGGA 6525
Db 6415 TTTCCGCTGAGCAGACATGCTGGAGAGGTATTTCTGTACTTCTCAAGGACACAGGGGA 6474
Qy 6526 AGGACGAGGACACCGAAGACGAGAGGAGGAGGTGGAGTGGACCCCGCCAGGCC 6585
Db 6475 AGGACGAGGACACCGAAGACGAGAGGAGGAGGTGGAGTGGACCCCGCCAGGCC 6534
Qy 6586 TGAGACACCCCAACCGGTTCAGCAGTTCCTCGATGACCTAGCAGTCCGAGACTGTGC 6645
Db 6535 TGAGACACCCCAACCGGTTCAGCAGTTCCTCGATGACCTAGCAGTCCGAGACTGTGC 6594
Qy 6646 TCTGAGCTCCCTCCCTCGCGGCGCGCGGAGGCGCTGGGAATGGCAAGGGCAAGGTA 6705
Db 6595 TCTGAGCTCCCTCCCTCGCGGCGCGCGGAGGCGCTGGGAATGGCAAGGGCAAGGTA 6654
Qy 6706 GAGTGGCTAGGAGCCCTGGACTCAGGCTGCGAGAGGGGCTGGTCCCTGGAGAAATAA 6765
Db 6655 GAGTGGCTAGGAGCCCTGGACTCAGGCTGCGAGAGGGGCTGGTCCCTGGAGAAATAA 6714
Qy 6766 GAGAGGCTGAGAGAGCCGCTGCTGTGTGAAAAA 6804
Db 6715 GAGAGGCTGAGAGAGCCGCTGCTGTGTGAAAAA 6753

RESULT 5

ABL57810 standard; DNA; 7795 BP.

XX ABL57810;

XX 05-AUG-2002 (first entry)

XX Human ABCA7 coding sequence.

XX Human; ABCA7; promoter; immunomodulatory; antiinflammatory; metabolic;
XX ATP-Binding Cassette; lipid metabolism disorder; immune response;
XX inflammation; gene therapy; gene; ds.

XX Homo sapiens.

XX WO200234903-A2.

XX 02-MAY-2002.

XX 17-OCT-2001; 2001WO-FR003219.

XX 24-OCT-2000; 2000FR-00013649.

XX 28-NOV-2000; 2000US-0253141P.

XX (AVET) AVENTIS PHARMA SA.

XX (NRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Denefle P, Rosier M, Prades C, Arnould-Reguigne I;

XX Osorio Y ForteauJ, Duverger N, Chimini G;

XX WPI; 2002-362799/39.

XX New promoter of the ABCA7 gene, useful for identifying modulators of

XX transcription and in gene therapy of e.g. disorders of lipid metabolism.

XX Disclosure; Page 124-125; 126pp; French.

XX The present invention relates to ABCA7 gene promoter sequences (ABC
CC stands for ATP-Binding Cassette), which are used to identify agents (A)
CC that modulate transcription of nucleic acids placed under control of the
CC promoter. (A) is potentially useful for treating or preventing defects in
CC lipid metabolism and defects in mechanisms involved in the immune
CC response and inflammation. The promoters can also be used in gene therapy
CC to control expression of therapeutic genes. Analysis of the promoter
CC sequences can be used diagnostically, particularly to identify subjects
CC at risk of lipid metabolism disorders. The present sequence is a coding
CC sequence for human ABCA7, used to illustrate the invention
XX
SQ Sequence 7795 BP; 1328 A; 2477 C; 2488 G; 1502 T; 0 U; 0 Other;

Query Match 98.6%; Score 6710; DB 6; Length 7795;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 6722; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 56 TGGGAGGCGCCCTGGAGAGCTTCCAGGAACCTCGCTGTGGGATAAAGGAATGAGTTTC 115
Db 1054 TGCTCGCTCTAATCAGAGCTTCCAGGAACCTCGCTGTGGGATAAAGGAATGAGTTTC 1113
Qy 116 AGAAAGGGGCGAGGAGTTGCCCGCAGCCGACCGCAGCTTTCAGCCCGACCGTTTCCT 175
Db 1114 AGAAAGGGGCGAGGAGTTGCCCGCAGCCGACCGCAGCTTTCAGCCCGACCGTTTCCT 1173
Qy 176 GACCTCTCTGCTCGCTGCCCGCCCGCAGTCTCACCATGGCCTTCTGGACACAGCTGATCT 235
Db 1174 GACCTCTCTGCTCGCTGCCCGCCCGCAGTCTCACCATGGCCTTCTGGACACAGCTGATCT 1233
Qy 236 GCTGCTCTCGAAGAAATTTCACTGTATCGCGGAGAGCAGCCGCTCCAGCTCTTGTCGAAAT 295
Db 1234 GCTGCTCTCGAAGAAATTTCACTGTATCGCGGAGAGCAGCCGCTCCAGCTCTTGTCGAAAT 1293
Qy 296 GCTGTGGCCTCTCTTCTCTTCTCATCTGCTGGTGTTCGCCCACTCCCAACCCGCCCT 355
Db 1294 GCTGTGGCCTCTCTTCTCTTCTCATCTGCTGGTGTTCGCCCACTCCCAACCCGCCCT 1353
Qy 356 GGAGCACATGATGCGCACTTCCCAAAAGCCACTGCCATCGGGGGGACCGTGGCCTG 415
Db 1354 GGAGCACATGATGCGCACTTCCCAAAAGCCACTGCCATCGGGGGGACCGTGGCCTG 1413
Qy 416 GCTCCAGGGTCTCATCTGTAATGTGAACAAACCTCTTTTCGCGAGCTGACACCGGCGGA 475
Db 1414 GCTCCAGGGTCTCATCTGTAATGTGAACAAACCTCTTTTCGCGAGCTGACACCGGCGGA 1473
Qy 476 GGAGCCCGGGCGCTGAGCAACTTCAACGACTCCCTGGTCTCCCGCTGCTAGCCGATGC 535
Db 1474 GGAGCCCGGGCGCTGAGCAACTTCAACGACTCCCTGGTCTCCCGCTGCTAGCCGATGC 1533
Qy 536 CGCACTGTCTGGGAGGGGCGAGTCCCAAGGAGCGTGGCTGGCTAGGGAGCTGAT 595
Db 1534 CGCACTGTCTGGGAGGGGCGAGTCCCAAGGAGCGTGGCTGGCTAGGGAGCTGAT 1593
Qy 596 CGCCACGCTGAGGGCTGACGAGCAGCGCCAGCGCTCAACCAACCAAGCAGTCCACT 655
Db 1594 CGCCACGCTGAGGGCTGACGAGCAGCGCCAGCGCTCAACCAACCAAGCAGTCCACT 1653
Qy 656 GGAAACCAACCATGCTGGATGTGCGGAGCTGCTGACGTCACTGCTGCGCACCGGAATCCCT 1715
Db 1654 GGAAACCAACCATGCTGGATGTGCGGAGCTGCTGACGTCACTGCTGCGCACCGGAATCCCT 1713
Qy 716 GGGTTGGCACTGGGCGCAAGCCGAGAGCCCTTGCAAGCTTGTGGAGGCCCTGAGGA 775
Db 1714 GGGTTGGCACTGGGCGCAAGCCGAGAGCCCTTGCAAGCTTGTGGAGGCCCTGAGGA 1773
Qy 776 CCTGGCCCGAGGAGCTCTGCGCTGCGCAGCTGTGGAGCTTCGCGGCACTGCTGCAGAG 835
Db 1774 CCTGGCCCGAGGAGCTCTGCGCTGCGCAGCTGTGGAGCTTCGCGGCACTGCTGCAGAG 1833
Qy 836 ACCCCGAGGAGACCGAGCGGCCCTCGAGTTGCTGTGTCAGAGGCCCTCTGAGTGTGAGGG 895
Db 1834 ACCCCGAGGAGACCGAGCGGCCCTCGAGTTGCTGTGTCAGAGGCCCTCTGAGTGTGAGGG 1893

QY 5276 TATCTTCCCCCACTTCTGCTTGGGCGGGGGCTCAATTGACATGTTGCGGAAACAGGCCAT 5335
DB 6274 TATCTTCCCCCACTTCTGCTTGGGCGGGGGCTCAATTGACATGTTGCGGAAACAGGCCAT 6333
QY 5336 GGTCTGATGCTTTGAGCGCTTGGGAGACAGGCAGTTCCAGTACACCCCTGCGTGGAGGT 5395
DB 6334 GGTCTGATGCTTTGAGCGCTTGGGAGACAGGCAGTTCCAGTACACCCCTGCGTGGAGGT 6393
QY 5396 GGTCTGATGCTTTGAGCGCTTGGGAGACAGGCAGTTCCAGTACACCCCTGCGTGGAGGT 5455
DB 6394 GGTCTGATGCTTTGAGCGCTTGGGAGACAGGCAGTTCCAGTACACCCCTGCGTGGAGGT 6453
QY 5456 ACTGCTGAGCAACCGAAGCAACTCTCTGCAAGCCAGCCAGCGGTGAGGTCTCTGCCACTCCT 5515
DB 6454 ACTGCTGAGCAACCGAAGCAACTCTCTGCAAGCCAGCCAGCGGTGAGGTCTCTGCCACTCCT 6513
QY 5516 GGGAGAGGAGCAGGAGTGTAGCCGTTGACCGGAGCGGGTGTCCAGAGAGCCACCA 5575
DB 6514 GGGAGAGGAGCAGGAGTGTAGCCGTTGACCGGAGCGGGTGTCCAGAGAGCCACCA 6573
QY 5576 GGGGATGCTGTTGGTGTGAGAACTTGACCAAGGTATACCGTGGGCGAGAGTCCACAGC 5635
DB 6574 GGGGATGCTGTTGGTGTGAGAACTTGACCAAGGTATACCGTGGGCGAGAGTCCACAGC 6633
QY 5636 TGTGACCGCTTGTCCCTGGGATTTCCCTCTGTTGAGTGTGTTTGGGCTGCTGGGTGTGAA 5695
DB 6634 TGTGACCGCTTGTCCCTGGGATTTCCCTCTGTTGAGTGTGTTTGGGCTGCTGGGTGTGAA 6693
QY 5696 TGGAGCAGGAGAGCTTCACTTTTCGATAGTGTGACGGGGACACATTTGGCAGCAGGG 5755
DB 6694 TGGAGCAGGAGAGCTTCACTTTTCGATAGTGTGACGGGGACACATTTGGCAGCAGGG 6753
QY 5756 CGAGCTGTGCTGGCAGCCACAGCGTGGCGCGGAAACCCAGTGTGCGCACCTCAGCAT 5815
DB 6754 CGAGCTGTGCTGGCAGCCACAGCGTGGCGCGGAAACCCAGTGTGCGCACCTCAGCAT 6813
QY 5816 GGGATACTGCTCAATCCGATGTCATCTTTGAGTGTGACGGGCGCGAGCACCTTGA 5875
DB 6814 GGGATACTGCTCAATCCGATGTCATCTTTGAGTGTGACGGGCGCGAGCACCTTGA 6873
QY 5876 CTTGCTGCGGCTGCGCGTGTCCCGAGGCGCCAGTGTGCGCACCGCTGCTGAG 5935
DB 6874 CTTGCTGCGGCTGCGCGTGTCCCGAGGCGCCAGTGTGCGCACCGCTGCTGAG 6933
QY 5936 CTTGCGGCTGCTGGACTCTCATGTAGCAGACCGGCTGCGAGCACCTACAGCGAGG 5995
DB 6934 CTTGCGGCTGCTGGACTCTCATGTAGCAGACCGGCTGCGAGCACCTACAGCGAGG 6993
QY 5996 GAACAAACGCAAGCTGGCGACGGCCCTGGCTGTGGGAGCCAGCCGCTGCTTCT 6055
DB 6994 GAACAAACGCAAGCTGGCGACGGCCCTGGCTGTGGGAGCCAGCCGCTGCTTCT 7053
QY 6056 GGAAGAGCCGACACAGGAGTGAACCCAGCGCGCGGCTTCTTTTGAACACGCTTTT 6115
DB 7054 GGAAGAGCCGACACAGGAGTGAACCCAGCGCGCGGCTTCTTTTGAACACGCTTTT 7113
QY 6116 GGGCTGTGCGGAGGCGGCTTCACTGATGCTCCTCCATAGCATGAGGAGTGA 6175
DB 7114 GGGCTGTGCGGAGGCGGCTTCACTGATGCTCCTCCATAGCATGAGGAGTGA 7173
QY 6176 AGCGCTGTGCTGCGGCTTAGCATCATGTTGAATGGGCGGTTCCGCTGCGTGGCAGCC 6235
DB 7174 AGCGCTGTGCTGCGGCTTAGCATCATGTTGAATGGGCGGTTCCGCTGCGTGGCAGCC 7233
QY 6236 GCAACATCTAAGGCGAGATTTGCGGCGGTTCAACACTGACCTGCGGTTGCCCGCG 6295
DB 7234 GCAACATCTAAGGCGAGATTTGCGGCGGTTCAACACTGACCTGCGGTTGCCCGCG 7293
QY 6296 AAGTCCCGAGCGGAGGCTTGTGGCGGCGAGTTCCTTGGTGGAGCTGCGGA 6355
DB 7294 AAGTCCCGAGCGGAGGCTTGTGGCGGCGAGTTCCTTGGTGGAGCTGCGGA 7353

QY 6356 GGCACATGAGAGCGCGCTTCCAGCTTCCAGCTGCGCGCGGAGGCGCTGCGCCTTGGCGG 6415
DB 7354 GGCACATGAGAGCGCGCTTCCAGCTTCCAGCTGCGCGCGGAGGCGCTGCGCCTTGGCGG 7413
QY 6416 CGTCTTTGGAGAGCTTGGCGGTGCACGGCGCAGAGCACGGCTGGAGGACTTTTCCGTGAG 6475
DB 7414 CGTCTTTGGAGAGCTTGGCGGTGCACGGCGCAGAGCACGGCTGGAGGACTTTTCCGTGAG 7473
QY 6476 CCAGACGATGCTGGAGGAGTATTCTTGTACTTCTTCCAAAGGACACAGGGGAGGAGGAGGA 6535
DB 7474 CCAGACGATGCTGGAGGAGTATTCTTGTACTTCTTCCAAAGGACACAGGGGAGGAGGAGGA 7533
QY 6536 CACCAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCAGGCTTGCAGCACCC 6595
DB 7534 CACCAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCAGGCTTGCAGCACCC 7593
QY 6596 CAAAGCGGTGAGCCAGTTCTTCGATGATCCCTAGCACTGCCAGACTGTGCTTGAAGCTC 6655
DB 7594 CAAAGCGGTGAGCCAGTTCTTCGATGATCCCTAGCACTGCCAGACTGTGCTTGAAGCTC 7653
QY 6656 CCTCCCTTGGCGGCGCGGAGGCGCTTGGAAATGGCAAGGGCAAGGTAGAGTGCCTAG 6715
DB 7654 CCTCCCTTGGCGGCGCGGAGGCGCTTGGAAATGGCAAGGGCAAGGTAGAGTGCCTAG 7713
QY 6716 GAGCCCTGGACTCAGGCTGGCAGAGGGGCTGCTGCTTGGAGAAATAAGAGAGGCTG 6775
DB 7714 GAGCCCTGGACTCAGGCTGGCAGAGGGGCTGCTGCTTGGAGAAATAAGAGAGGCTG 7773
QY 6776 GAGAGAGCCGTGCTGTGTA 6797
DB 7774 GAGAGAGCCGTGCTGTGTA 7795

RESULT 6
AAS08706
ID AAS08706 standard; cDNA; 6522 BP.
XX
AC AAS08706;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human PD-ATP-binding cassette (PD-ABC) form #1 cDNA.
XX
KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus; ss;
KW peripheral blood leukocyte; bone marrow; lymph node; dyalipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..6441
FT /+tag= a
FT /product= "Human PD-ABC form #1"
FT
FN MO200153490-Al.
XX
PD 26-JUL-2001.
XX
XX 23-JAN-2001; 2001WO-US002191.
XX
PR 24-JAN-2000; 2000US-0177899P.
XX
PR 30-JUN-2000; 2000US-0215405P.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Johns MA, Tafuri SR, Wang M;
XX
DR WPI; 2001-442259/47.
DR P-PSDB; AAU04483.

Db 1801 GGGCCCTTCTGCTCAGCGCGCGCTGCTGTTCTGTTGCTCAAGCTGGGGGACATCCTC 1860
Qy 2070 CCTACAGCCACCCGGGGTGTCTTCTGTTCTTGGCAGCCTTCCGGTGGCCACCGTG 2129
Db 1861 CCTCGACCCACCGGGGTGTCTTCTGTTCTTGGCAGCCTTCCGGTGGCCACCGTG 1920
Qy 2130 ACCGAGACTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCTGGCGC 2189
Db 1921 ACCGAGACTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCTGGCGC 1980
Qy 2190 GGCCTGGCTTCTTCTCTTCTACCTGCGCTACGTTGCTGTGTGGCTTGGCGGACCGG 2249
Db 1981 GGCCTGGCTTCTTCTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Qy 2250 CTGCGCGGGTGGCGGGTGGCGCGAGCCTGCTGTGCGCGCTGGCTTGGCTTGGC 2309
Db 2041 CTGCGCGGGTGGCGGGTGGCGCGAGCCTGCTGTGCGCGCTGGCTTGGCTTGGC 2100
Qy 2310 TCGGAGAGCTGCTCTGCTGAGAGAGCAGGGCGAGGGCGCGAGTGCGCAACCTGGCG 2369
Db 2101 TCGGAGAGCTGCTCTGCTGAGAGAGCAGGGCGAGGGCGCGAGTGCGCAACCTGGCG 2160
Qy 2370 ACCGGCTTACGGCAGAGCTTCTAGCCTGCGCCAGAGTCTTGGCTTCTGCTGCTGAC 2429
Db 2161 ACCGGCTTACGGCAGAGCTTCTAGCCTTGGCCAGAGTCTTGGCTTCTGCTGCTGAC 2220
Qy 2430 GCGGCTTCTACGGCTCGCCACCTGGTACCTGGAAGCTGTGTGCCAGGCGAGTACGG 2489
Db 2221 GCGGCTTCTACGGCTCGCCACCTGGTACCTGGAAGCTGTGTGCCAGGCGAGTACGG 2280
Qy 2490 ATCCCTGAACATGGAATTTTCTTTTTCGGAGAGCTACTGGTGGGACCTCGGCCCCC 2549
Db 2281 ATCCCTGAACATGGAATTTTCTTTTTCGGAGAGCTACTGGTGGGACCTCGGCCCCC 2340
Qy 2550 AAGAGTCAGCCCTTGGCCACCCCGCTGAGCCAGCCAGAGTGTCTGTAGAGAGCAGC 2609
Db 2341 AAGAGTCAGCCCTTGGCCACCCCGCTGAGCCAGCCAGAGTGTCTGTAGAGAGCAGC 2400
Qy 2610 CCGGCTTCTGCTGGGCTTCTGCTGCGAGCTTGGAGAGAGCTTCTTCTGGAAGCCG 2669
Db 2401 CCGGCTTCTGCTGGGCTTCTGCTGCGAGCTTGGAGAGAGCTTCTTCTGGAAGCCG 2460
Qy 2670 CAGCAGCCCTGCGGGGCTCAGCTGCACTTCTACAGGGCCACATCACCGCTTCTG 2729
Db 2461 CAGCAGCCCTGCGGGGCTCAGCTGCACTTCTACAGGGCCACATCACCGCTTCTG 2520
Qy 2730 GGCACAGGGGCGGCAAGACCACTCTGTCTTCTGAGTGGCTCTTCCACCC 2789
Db 2521 GGCACAGGGGCGGCAAGACCACTCTGTCTTCTGAGTGGCTCTTCCACCC 2580
Qy 2790 AGTGTGGCTTCTGCTTCTATCTGGGCGACGAGCTCGCTCCAGCATGGCGCCATCCG 2849
Db 2581 AGTGTGGCTTCTGCTTCTATCTGGGCGACGAGCTCGCTCCAGCATGGCGCCATCCG 2640
Qy 2850 CCCCACCTGGGCTGTCTCTCAGTACAACTGCTGTTTGAATGCTGACCGTGGAGAG 2909
Db 2641 CCCCACCTGGGCTGTCTCTCAGTACAACTGCTGTTTGAATGCTGACCGTGGAGAG 2700
Qy 2910 CAGCTCTGGTCTATGGCGGCTGAAGGCTGAGTGGCGCTGAGTGGGCGCCGAGAG 2969
Db 2701 CAGCTCTGGTCTATGGCGGCTGAAGGCTGAGTGGCGCTGAGTGGGCGCCGAGAG 2760
Qy 2970 GACGCTGTGTCAGGATGTGGGCTGTCTCCAGCAGAGTGTGACAGCTCGCCACCTC 3029
Db 2761 GACGCTGTGTCAGGATGTGGGCTGTCTCCAGCAGAGTGTGACAGCTCGCCACCTC 2820
Qy 3030 TCTGTGGGATGCAACGGAAGCTGTCTGCTGGCCATTGCTTTTGGGGCGGCTCCCAAGT 3089
Db 2821 TCTGTGGGATGCAACGGAAGCTGTCTGCTGGCCATTGCTTTTGGGGCGGCTCCCAAGT 2880
Qy 3090 GTTATCTGGAAGAGCTTACGGCTGCGTGATCTCTTCCCGCGGCTATTGGGAG 3149
Db 2881 GTTATCTGGAAGAGCTTACGGCTGCGTGATCTCTTCCCGCGGCTATTGGGAG 2940

Qy 3150 CTGCTGCTCAATACCGAGAAAGGTGCGACGCTGATCTCTCCACCCACCTCGATGAG 3209
Db 2941 CTGCTGCTCAATACCGAGAAAGGTGCGACGCTGATCTCTCCACCCACCTCGATGAG 3000
Qy 3210 GCAGAGCTGCTGGGAGACCGTGTGCTGTGGCAGGTGGCGGCTGTGCTGTGGC 3269
Db 3001 GCAGAGCTGCTGGGAGACCGTGTGCTGTGGCAGGTGGCGGCTGTGCTGTGGC 3060
Qy 3270 TCCCACTTCTTCTGCGCGCTCACCTGGCTTCCGGCTACTACTGACCTGTGTAAGGC 3329
Db 3061 TCCCACTTCTTCTGCGCGCTCACCTGGCTTCCGGCTACTACTGACCTGTGTAAGGC 3120
Qy 3330 CGCTGCTTCTGACCAACCAATGAGAAGCTGACACTGACATGAGAGGCGAGTGTGGAC 3389
Db 3121 CGCTGCTTCTGACCAACCAATGAGAAGCTGACACTGACATGAGAGGCGAGTGTGGAC 3180
Qy 3390 AGGAGGAAAGAAATGGCAGCGGACGAGAGTGGCAGCTCTCCTCAGTGTGCTG 3449
Db 3181 AGGAGGAAAGAAATGGCAGCGGACGAGAGTGGCAGCTCTCCTCAGTGTGCTG 3240
Qy 3450 CTGCTACAGCACTGGGTGCGCGGCGACGCTGTGGAGAGCTGCCACACGAGCTGTG 3509
Db 3241 CTGCTACAGCACTGGGTGCGCGGCGACGCTGTGGAGAGCTGCCACACGAGCTGTG 3300
Qy 3510 CTGCTGCTGCTTACACGGGTGCCATGACGGCAGCTTTCGCCACACTCTTCCGAGAGCTA 3569
Db 3301 CTGCTGCTGCTTACACGGGTGCCATGACGGCAGCTTTCGCCACACTCTTCCGAGAGCTA 3360
Qy 3570 GACACGGCTGCGGAGCTGAGGCTCCTGCTACGGGATCTCGACACGAGCTCGAG 3629
Db 3361 GACACGGCTGCGGAGCTGAGGCTCCTGCTACGGGATCTCGACACGAGCTCGAG 3420
Qy 3630 GAGATCTCTTGAAGTGTGGAGAGTGTCTCGGACACAGATATGGAGATGGCAGC 3689
Db 3421 GAGATCTCTTGAAGTGTGGAGAGTGTCTCGGACACAGATATGGAGATGGCAGC 3480
Qy 3690 TCGCGGCGAGCACTTATGACAGGCACTTCTGCTTCTAGACGTAACCTCGGCTCAAGATG 3749
Db 3481 TCGCGGCGAGCACTTATGACAGGCACTTCTGCTTCTAGACGTAACCTCGGCTCAAGATG 3540
Qy 3750 CCGCACAGGAGACAGCGCTGGAGAACGGGAAACAGCTGGGTGAGCCGACAGACTGAC 3809
Db 3541 CCGCACAGGAGACAGCGCTGGAGAACGGGAAACAGCTGGGTGAGCCGACAGACTGAC 3600
Qy 3810 CAGGCTCTGGGCGAGCGCTGGCGGCTGACGGCTGAGGCTGAGGCTGAGCCGAGAG 3869
Db 3601 CAGGCTCTGGGCGAGCGCTGGCGGCTGAGGCTGAGGCTGAGGCTGAGCCGAGAG 3660
Qy 3870 CTCCAGGCTTCTTCTCAAGCGCTTCTGCTTCTGCTGCGCGCGCAGCGCGGCTGCTTC 3929
Db 3661 CTCCAGGCTTCTTCTCAAGCGCTTCTGCTTCTGCTGCGCGCGCAGCGCGGCTGCTTC 3720
Qy 3930 GCCAGATCTGCTGCTGCTTCTGCTGCGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 3989
Db 3721 GCCAGATCTGCTGCTGCTTCTGCTGCGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 3780
Qy 3990 CTTCTTCTGGGCTTCTGCTGCGCTTCTGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 4049
Db 3781 CTTCTTCTGGGCTTCTGCTGCGCTTCTGCTGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 3840
Qy 4050 TCTTCTTCTAGTAGGACGCGCGCGGACCTTGGACGCTGCTGCGCTTCTGAGCGCTG 4109
Db 3841 TCTTCTTCTAGTAGGACGCGCGCGGACCTTGGACGCTGCTGCGCTTCTGAGCGCTG 3900
Qy 4110 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4169
Db 3901 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960
Qy 4170 CCAGAGTCTCTGCTGAGTGGCCCAAGGCTTGTGGCAGTGGCAACTGGACCCAGAGTCT 4229
Db 3961 CCAGAGTCTCTGCTGAGTGGCCCAAGGCTTGTGGCAGTGGCAACTGGACCCAGAGTCT 4020

QY 4230 CCATCCCGAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGTGCTCCCGAGTGCCTCCG 4289
DB |||||
QY 4021 CCATCCCGAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGTGCTCCCGAGTGCCTCCG 4080
DB |||||
QY 4290 GCTGCACTGGTGGTCCCCCTCCGCCCCAGGCAAGTGAACGGCTCTGGGGAAGTGGTTCAG 4349
DB |||||
QY 4081 GCTGCACTGGTGGTCCCCCTCCGCCCCAGGCAAGTGAACGGCTCTGGGGAAGTGGTTCAG 4140
DB |||||
QY 4350 AACCTGACAGCCGGAACCTGTCTGACTTCTGGTCAAGACCTACCCGCGCTGTGGTGGC 4409
DB |||||
QY 4141 AACCTGACAGCCGGAACCTGTCTGACTTCTGGTCAAGACCTACCCGCGCTGTGGTGGC 4200
DB |||||
QY 4410 CAGGGCTGAAGACTAAGAAAGTGGGTGAATGAGTCAAGTACAGAGCTTCTCGCTGGG 4469
DB |||||
QY 4201 CAGGGCTGAAGACTAAGAAAGTGGGTGAATGAGTCAAGTACAGAGCTTCTCGCTGGG 4260
DB |||||
QY 4470 GGGCGAGACCCAGGCTGCCCTCCGGCCAAAGTGTGGGCGCTCAGTGGAGGAGTGTGG 4529
DB |||||
QY 4261 GGGCGAGACCCAGGCTGCCCTCCGGCCAAAGTGTGGGCGCTCAGTGGAGGAGTGTGG 4320
DB |||||
QY 4530 GCGCTGCTGAGTCCCTGCCCTGCCGGGGCCCTCGAAGCTGTCTGAAACCTCACAGCC 4589
DB |||||
QY 4321 GCGCTGCTGAGTCCCTGCCCTGCCGGGGCCCTCGAAGCTGTCTGAAACCTCACAGCC 4380
DB |||||
QY 4590 TGGGCTCACAGCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAAGGCTGG 4649
DB |||||
QY 4381 TGGGCTCACAGCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAAGGCTGG 4440
DB |||||
QY 4650 CACTTCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCCCC 4709
DB |||||
QY 4441 CACTTCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCCCC 4500
DB |||||
QY 4710 CAGAGCCCGGCGCGCCAGCCCAAGCAATCAACACTCAACCAACCCCTTGAACCTCAC 4769
DB |||||
QY 4501 CAGAGCCCGGCGCGCCAGCCCAAGCAATCAACACTCAACCAACCCCTTGAACCTCAC 4560
DB |||||
QY 4770 AAGGAGCAGCTGTCTGAGGCTGACATGAGCTCTCGTGGAGCTCTCGTCTCATC 4829
DB |||||
QY 4561 AAGGAGCAGCTGTCTGAGGCTGACATGAGCTCTCGTGGAGCTCTCGTCTCATC 4620
DB |||||
QY 4830 TGTGTGCTCTTTCGCAATGCTCTTGTCCCGGCGAGCTTCACTCTTGTCTCATCTGAGGAG 4889
DB |||||
QY 4621 TGTGTGCTCTTTCGCAATGCTCTTGTCCCGGCGAGCTTCACTCTTGTCTCATCTGAGGAG 4680
DB |||||
QY 4890 CGAGTCAACCGAGCAAGCACTGCACTCATGAGGCTCTGTCGCCCAACCTCTACTGG 4949
DB |||||
QY 4681 CGAGTCAACCGAGCAAGCACTGCACTCATGAGGCTCTGTCGCCCAACCTCTACTGG 4740
DB |||||
QY 4950 CTTTGGCAACTTCTCTGGGACATGTGTAACCTCTTGTGGCCAGCATGCACTCGTGGTCTC 5009
DB |||||
QY 4741 CTTTGGCAACTTCTCTGGGACATGTGTAACCTCTTGTGGCCAGCATGCACTCGTGGTCTC 4800
DB |||||
QY 5010 ATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTCCCAACCTGCTCTCCTG 5069
DB |||||
QY 4801 ATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTCCCAACCTGCTCTCCTG 4860
DB |||||
QY 5070 CTGTTGCTACTACTGTATGCTGTGATCAACCGCTCATGATCAACCGCTCTCTTCTTC 5129
DB |||||
QY 4861 CTGTTGCTACTACTGTATGCTGTGATCAACCGCTCATGATCAACCGCTCTCTTCTTC 4920
DB |||||
QY 5130 TTCTCCGTGCCAGCAGACCTATGTGGTCTCACTGATCAACCTCTTATTTGGCATC 5189
DB |||||
QY 4921 TTCTCCGTGCCAGCAGACCTATGTGGTCTCACTGATCAACCTCTTATTTGGCATC 4980
DB |||||
QY 5190 AATGGAAGCAGTGGCCACTTGTGCTTGTGATCTTCTGATCAGAACTGCAAGAGGTG 5249
DB |||||
QY 4981 AATGGAAGCAGTGGCCACTTGTGCTTGTGATCTTCTGATCAGAACTGCAAGAGGTG 5040
DB |||||
QY 5250 AGCCGGATCTTGAACAGGTCTTCTTATCTTCCCCACCTTCTGCTTGGCCGGGGCTC 5309
DB |||||
QY 5041 AGCCGGATCTTGAACAGGTCTTCTTATCTTCCCCACCTTCTGCTTGGCCGGGGCTC 5100
DB |||||
QY 5310 ATTGACATGTGTGGGAACAGGCCATGGCTGTATGCCCTTTGAGCGCTTGGGAGACAGGCAG 5369

DB |||||
QY 5101 ATTGACATGTGTGGGAACAGGCCCATGGCTGTATGCCCTTTGAGCGCTTGGGAGACAGGCAG 5160
DB |||||
QY 5370 TTCCAGTCAACCTTGGCTGGGAGGTGGTCGGCAAGAACTCTTTGGCCATGGTGATACAG 5429
DB |||||
QY 5161 TTCCAGTCAACCTTGGCTGGGAGGTGGTCGGCAAGAACTCTTTGGCCATGGTGATACAG 5220
DB |||||
QY 5430 GGGCCCTCTTCTTCTCTTCTCACTACTGTGCACACCGAAGCAACTCTCTGCGCCACAG 5489
DB |||||
QY 5221 GGGCCCTCTTCTTCTCTTCTCACTACTGTGCACACCGAAGCAACTCTCTTGGCCACAG 5280
DB |||||
QY 5490 CCAGAGGTGAGGTCTCTGCACTCTCTGGGAGAGGAGAGAGATGTAGCCCTGGAACGG 5549
DB |||||
QY 5281 CCAGAGGTGAGGTCTCTGCACTCTCTGGGAGAGGAGAGAGATGTAGCCCTGGAACGG 5340
DB |||||
QY 5550 GAGCGGGTGGTCCAGAGGCCACCCAGGGGGATGTGTGGTGTGAGGAACTTGAACCAAG 5609
DB |||||
QY 5341 GAGCGGGTGGTCCAGAGGCCACCCAGGGGGATGTGTGGTGTGAGGAACTTGAACCAAG 5400
DB |||||
QY 5610 GTATACCGGTGGGAGAGGATGCCAGCTGTGACCGCTTGTGCTGGGGATTTCCCTCTGT 5669
DB |||||
QY 5401 GTATACCGGTGGGAGAGGATGCCAGCTGTGACCGCTTGTGCTGGGGATTTCCCTCTGT 5460
DB |||||
QY 5670 GAGTGTGTTGGGCTGTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTTCGATGGTG 5729
DB |||||
QY 5461 GAGTGTGTTGGGCTGTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTTCGATGGTG 5520
DB |||||
QY 5730 ACGGGGGAACAATTTGCCAGAGGGCGAGGCTGTCTGGCAGGGCCACAGCGTGGGCCGG 5789
DB |||||
QY 5521 ACGGGGGAACAATTTGCCAGAGGGCGAGGCTGTCTGGCAGGGCCACAGCGTGGGCCGG 5580
DB |||||
QY 5790 GAACCCAGTGTGGCGCACCTCAGCATGGGATCTGCCCTCNACTCCATGCCATCTTTGAG 5849
DB |||||
QY 5581 GAACCCAGTGTGGCGCACCTCAGCATGGGATCTGCCCTCAATCCGATGCCATCTTTGAG 5640
DB |||||
QY 5850 CTGCTGACGGGCGCGCAGCACTTGGAGCTGTCTGGCGGCTGTGGCGGCTGTCCGGAGGCC 5909
DB |||||
QY 5641 CTGCTGACGGGCGCGCAGCACTTGGAGCTGTCTGGCGGCTGTGGCGGCTGTCCGGAGGCC 5700
DB |||||
QY 5910 CAGTGTGCCAGACCCGCTGCTCAGGCTGTGGCGGCTGTGGGACTCTCATGTGTAAGGAGAC 5969
DB |||||
QY 5701 CAGTGTGCCAGACCCGCTGCTCGGCGCTGGCGGCTGTGGGACTCTCATGTGTAAGGAGAC 5760
DB |||||
QY 5970 CGGCTGTGAGGCACTTACAGCGGAGGGAACAACGCAAGCTGCGCAGCGGCTGTGGCGCTG 6029
DB |||||
QY 5761 CGGCTGTGAGGCACTTACAGCGGAGGGAACAACGCAAGCTGCGCAGCGGCTGTGGCGCTG 5820
DB |||||
QY 6030 GTTGGGGAACCCAGCCGTGTGTTCTTGGACGAGCCGACCAAGCCATGGACCCAGCGCG 6089
DB |||||
QY 5821 GTTGGGGAACCCAGCCGTGTGTTCTTGGACGAGCCGACCAAGCCATGGACCCAGCGCG 5880
DB |||||
QY 6090 CGGCGCTTCTTTGGAAACAGCCTTTTGGCGGCTGTGGGAGGGCGGCTTCAGTGATGCTC 6149
DB |||||
QY 5881 CGGCGCTTCTTTGGAAACAGCCTTTTGGCGGCTGTGGGAGGGCGGCTTCAGTGATGCTC 5940
DB |||||
QY 6150 ACCTCCATAGCATGAGGAGGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGTAAT 6209
DB |||||
QY 5941 ACCTCCATAGCATGAGGAGGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGTAAT 6000
DB |||||
QY 6210 GGGCGGTTCGCGCTGCTGGGAGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCAAC 6269
DB |||||
QY 6001 GGGCGGTTCGCGCTGCTGGGAGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCAAC 6060
DB |||||
QY 6270 ACATGACCTGCGGTGCGCGCCGCAAGGTCCAGCCGCGAGCGGCTTGTGGCGGCGC 6329
DB |||||
QY 6061 ACATGACCTGCGGTGCGCGCCGCAAGGTCCAGCCGCGAGCGGCTTGTGGCGGCGC 6120
DB |||||
QY 6330 GAGTTCCTGGGTGCGAGCTGCGGAGGCAACATGAGAGGCGGCTGCGCTTCCAGCTGCG 6389
DB |||||
QY 6121 GAGTTCCTGGGTGCGAGCTGCGGAGGCAACATGAGAGGCGGCTGCGCTTCCAGCTGCG 6180
DB |||||
QY 6390 CCGGAGGGCGCTGCGCCCTTGGCGCGCTTGTGGAGAGCTGGCGGTGCAACGCGGAGAG 6449

QY 870 TCAGAGCCCTCTCAGTGT CAGGGACCTAGACGACAGTGGGCCCTCCCTCAACTGG 929
Db 655 TCAGAGCCCTCTTGAGTGT CAGGGACCTAGACGACAGTGGGCCCTCCCTCAACTGG 714
QY 930 TACGAGCTAGTACCTGATGGAGCTGTGTGGGACGAGCCAGATCGCGCTTCCAGAC 989
Db 715 TACGAGCTAGTACCTGATGGAGCTGTGTGGGACGAGCCAGATCGCGCTTCCAGAC 774
QY 990 AGCAGCTGAGCCCGCTCTCTCGAGCTGATTTGGAGCCCTGACAGCCAGCCGTGTCC 1049
Db 775 AGCAGCTGAGCCCGCTCTCTCGAGCTGATTTGGAGCCCTGACAGCCAGCCGTGTCC 834
QY 1050 CGCTGTCTTGGAGAGCCCTGAAGCTTCTGATCTCTGGGAAGCTACTTTTGCACAGAT 1109
Db 835 CGCTGTCTTGGAGAGCCCTGAAGCTTCTGATCTCTGGGAAGCTACTTTTGCACAGAT 894
QY 1110 ACACCTTTTACCCGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAGCTCACCCCTG 1169
Db 895 ACACCTTTTACCCGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAGCTCACCCCTG 954
QY 1170 CTGAGGATGTCTGGGAGGTGTGGAGATCTGGGACCCCGGATCTTACCTTCATGAAC 1229
Db 955 CTGAGGATGTCTGGGAGGTGTGGGAGATCTGGGACCCCGGATCTTACCTTCATGAAC 1014
QY 1230 GACAGTTTCCAAATGTGGCCATGTCTGACGGCTCTCTGCAGATGAGGATGAAGGAAGG 1289
Db 1015 GACAGTTTCCAAATGTGGCCATGTCTGACGGCTCTCTGCAGATGAGGATGAAGGAAGG 1074
QY 1290 CAGCCAGACTGTGAGGCGGGACCAATGAGGCGCTTCGATCTCTTTTGGACCCCTGGG 1349
Db 1075 CAGCCAGACTGTGAGGCGGGACCAATGAGGCGCTTCGATCTCTTTTGGACCCCTGGG 1134
QY 1350 AGCGTGGCTTACAGCTGGAGGAGGACACAGCTGATGTGGGGACCTGTGTGGGACGCTG 1409
Db 1135 AGCGTGGCTTACAGCTGGAGGAGGACACAGCTGATGTGGGGACCTGTGTGGGACGCTG 1194
QY 1410 GGCAGGTGACGAGTGCCTGTCTTGGACAACTGGAGCGGACCCCTCAGAGGAGGCC 1469
Db 1195 GGCAGGTGACGAGTGCCTGTCTTGGACAACTGGAGCGGACCCCTCAGAGGAGGCC 1254
QY 1470 CTGGTGTGCGGGCCCTTGCAACTGCTCGGGAACATCGATTTCTGGGCGGCGTCTTC 1529
Db 1255 CTGGTGTGCGGGCCCTTGCAACTGCTCGGGAACATCGATTTCTGGGCGGCGTCTTC 1314
QY 1530 TTGGGACCTGAGACTCTTACAGACCCACAGAGACCCACAGACCTGGGCGGCGG 1589
Db 1315 TTGGGACCTGAGACTCTTACAGACCCACAGAGACCCACAGACCTGGGCGGCGG 1374
QY 1590 CAGTGGCATCAAAATCCGATGGACATTTGACGTGTACAGAGGACCAATAAGATCAGG 1649
Db 1375 CAGTGGCATCAAAATCCGATGGACATTTGACGTGTACAGAGGACCAATAAGATCAGG 1434
QY 1650 GACAGGTTTTGGACCTTGCCCGGACCGGACCCCTGACCCGACCTGCGCTACGTGTGG 1709
Db 1435 GACAGGTTTTGGACCTTGCCCGGACCGGACCCCTGACCCGACCTGCGCTACGTGTGG 1494
QY 1710 GCGGCTTCTGTACTGTCAAGACTGTGTGGAGGTGACCGCTGCGGTGTCTCAGGGC 1769
Db 1495 GCGGCTTCTGTACTGTCAAGACTGTGTGGAGGTGACCGCTGCGGTGTCTCAGGGC 1554
QY 1770 GCCAACCCCGGCGGCTCTACCTGACAGATGCCCTATCCGTGCTATGTGGAGCAG 1829
Db 1555 GCCAACCCCGGCGGCTCTACCTGACAGATGCCCTATCCGTGCTATGTGGAGCAG 1614
QY 1830 GTGTTCCTGCTGTGTGAGCCGGTGCCTCGCTCTTCTGACGCTGGCTGGATCTAC 1889
Db 1615 GTGTTCCTGCTGTGTGAGCCGGTGCCTCGCTCTTCTGACGCTGGCTGGATCTAC 1674
QY 1890 TCCGTGACATGACAGTGAAGGCGGTGTGGGAGGAGGAGGAGCGGCTGCGGGACAC 1949
Db 1675 TCCGTGACATGACAGTGAAGGCGGTGTGGGAGGAGGAGGAGCGGCTGCGGGACAC 1734
QY 1950 ATGCGGCCATGGGCTCAGCGCGGCTGCTCTGGCTAGGCTGTTCTCTCAGCTGCCTC 2009

Db 1735 ATGCGGCCATGGGCTCAGCGCGGCTGCTCTGGCTAGGCTGTTCTCAGCTGCCTC 1794
QY 2010 GGGCCCTTCTCTGCTCAGCGCGGCTGCTGTTCTGTTCTGTTCTCAGCTGGGGACATCCTC 2069
Db 1795 GGGCCCTTCTCTGCTCAGCGCGGCTGCTGTTCTGTTCTCAGCTGGGGACATCCTC 1854
QY 2070 CCTACAGACACCCGCGGCTGTTCTCTGTTCTTCTGCGAGCCTTCGCGGTGGCCAGGTCG 2129
Db 1855 CCTACAGACACCCGCGGCTGTTCTCTGTTCTTCTGCGAGCCTTCGCGGTGGCCAGGTCG 1914
QY 2130 ACCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAAACCTGGCTGCGGCTCGGCG 2189
Db 1915 ACCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCCAAACCTGCTGCGGCTCGGCG 1974
QY 2190 GGGCTGGCTACTTCTCTCTCTACCTGCGCTGCTGTGTGTGTGTGTGGCGGAGCCGG 2249
Db 1975 GGGCTGGCTACTTCTCTCTCTACCTGCGCTGCTGTGTGTGTGTGTGGCGGAGCCGG 2034
QY 2250 CTGCGCGGCTGGCGGCTGGCGGAGCTGCTGTCTGCGCGTGGCTTCTGGCTTCGCG 2309
Db 2035 CTGCGCGGCTGGCGGCTGGCGGAGCTGCTGTCTGCGCGTGGCTTCTGGCTTCGCG 2094
QY 2310 TGGAGAGCTTGGCTTCTGCTGGAGGAGGAGGCGGCGGCGGCGGAGCTGGCAACAAGTGGG 2369
Db 2095 TGGAGAGCTTGGCTTCTGCTGGAGGAGGAGGCGGCGGCGGCGGAGCTGGCAACAAGTGGG 2154
QY 2370 ACCCGGCTACGGGAGAGCTTTCAGCTGGCGGAGCTCTCTGGCTTCTGGCTTCGCG 2429
Db 2155 ACCCGGCTACGGGAGAGCTTTCAGCTGGCGGAGCTCTCTGGCTTCTGGCTTCGCG 2214
QY 2430 GGGCGCTCTACGGCTTCGCCACCTGTAAGCTGTGTGCGGCGGAGTGTGAGGAGTACGGG 2489
Db 2215 GGGCGCTCTACGGCTTCGCCACCTGTAAGCTGTGTGCGGCGGAGTGTGAGGAGTACGGG 2274
QY 2490 ATCCCTGAACCAATGGAATTTTCTTTTCGAGGAGCTACTGTGTGCGGAGCTTCGCGCCCC 2549
Db 2275 ATCCCTGAACCAATGGAATTTTCTTTTCGAGGAGCTACTGTGTGCGGAGCTTCGCGCCCC 2334
QY 2550 AGAGTCCAGCGCTTCGCCACCCCGCTGGACCCAAAGTGTGTGAGGAGGAGCAGG 2609
Db 2335 AGAGTCCAGCGCTTCGCCACCCCGCTGGACCCAAAGTGTGTGAGGAGGAGCAGG 2394
QY 2610 CCGGCTCAGTCTCGCGCTATCCGTTTCGAGCTGAGAGGCTTCTCTGGAAGCCG 2669
Db 2395 CCGGCTCAGTCTCGCGCTATCCGTTTCGAGCTGAGAGGCTTCTCTGGAAGCCG 2454
QY 2670 CAGCCAGCCCTGCGGGGCTCAGCTTGTCTTACAGGGCCATCATCAGCGCTTCTCTG 2729
Db 2455 CAGCCAGCCCTGCGGGGCTCAGCTTGTCTTACAGGGCCATCATCAGCGCTTCTCTG 2514
QY 2730 GGCACAAAGGGGCGGGAAGACCAACCTGTGTCTTGTGAGTGGCTTCTTCCACCC 2789
Db 2515 GGCACAAAGGGGCGGGAAGACCAACCTGTGTCTTGTGAGTGGCTTCTTCCACCC 2574
QY 2790 AGTGTGGCTTCTGCTTCTTCTGCGGCGGAGCTGCGCTCCAGCATGCGCGGCTCCGG 2849
Db 2575 AGTGTGGCTTCTGCTTCTTCTGCGGCGGAGCTGCGCTCCAGCATGCGCGGCTCCGG 2634
QY 2850 CCCACCTTGGGCTTGTCTCTCAGTACAAAGTGTCTTTTGACATGCTGACCGTGGAGCAG 2909
Db 2635 CCCACCTTGGGCTTGTCTCTCAGTACAAAGTGTCTTTTGACATGCTGACCGTGGAGCAG 2694
QY 2910 CAGCTGTGGTCTTATGGGCGGTGAAGGCTGTGAGTGGCGCTGTAGTGGGCCCCGAGCAG 2969
Db 2695 CAGCTGTGGTCTTATGGGCGGTGAAGGCTGTGAGTGGCGCTGTAGTGGGCCCCGAGCAG 2754
QY 2970 GACGCTGTGAGGATGTGGGCTGTCTCCAAAGCAGAGTGTGAGACTGCCACCTC 3029
Db 2755 GACGCTGTGAGGATGTGGGCTGTCTCCAAAGCAGAGTGTGAGACTGCCACCTC 2814
QY 3030 TCTGTGGATGCAACGGAAGCTTCTGCTGGCCATTCCTTTGTGGGCGGCTCCCAAGTT 3089

Db 2815 TCTGTGGATGCAACGGAAGCTGTCGTGGCCATTGCTTGTGGCGGCTCCCAAGTT 2874
Qy 3090 GTTATCTTGGACAGCCTACCGCTGGCGTGGATCTGCTTCCGCGCGGGTATTTGGGAG 3149
Db 2875 GTTATCTTGGACAGCCTACCGCTGGCGTGGATCTGCTTCCGCGCGGGTATTTGGGAG 2934
Qy 3150 GTGCTGCTCAATACCGAAGGTCGACGCTGATCTCTCACCAACCACTGTGATGAG 3209
Db 2935 GTGCTGCTCAATACCGAAGGTCGACGCTGATCTCTCACCAACCACTGTGATGAG 2994
Qy 3210 GCAGAGCTGCTGGGAGACGCTGTGCTGTGTGGCAGGTGGCGCTTGTGCTGTGCTGGC 3269
Db 2995 GCAGAGCTGCTGGGAGACGCTGTGCTGTGTGGCAGGTGGCGCTTGTGCTGTGCTGGC 3054
Qy 3270 TCCCACTCTTCTCGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC 3329
Db 3055 TCCCACTCTTCTCGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC 3114
Qy 3330 GCCTGCCCTTGACCAACCAATGAGAAGGCTGACCTGACATGAGAGGGCAGTGTGACACC 3389
Db 3115 GCCTGCCCTTGACCAACCAATGAGAAGGCTGACCTGACATGAGAGGGCAGTGTGACACC 3174
Qy 3390 AGGCAGGAAAGAAAGAAATGGCAGCCAGGCGACAGAGTGGCACTCTCAGCTGTGCCC 3449
Db 3175 AGGCAGGAAAGAAAGAAATGGCAGCCAGGCGACAGAGTGGCACTCTCAGCTGTGCCC 3234
Qy 3450 CTGTGTACAGCACTGGGTGCCCGGGGACCGGCTGTGTGGAGAGCTGCCACACGAGCTGGTG 3509
Db 3235 CTGTGTACAGCACTGGGTGCCCGGGGACCGGCTGTGTGGAGAGCTGCCACACGAGCTGGTG 3294
Qy 3510 CTGTGTGTCCTTACACGGGTGCCCAATGACGGCAGCTTTCGCACTCTTCCGAGAGCTA 3569
Db 3295 CTGTGTGTCCTTACACGGGTGCCCAATGACGGCAGCTTTCGCACTCTTCCGAGAGCTA 3354
Qy 3570 GACACGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACGAGCTCGAG 3629
Db 3355 GACACGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACGAGCTCGAG 3414
Qy 3630 GAGATCTTCTGAAGGTGGTGGAGAGTGTGTGGGACACAGATATGAGAGTGGCAGC 3689
Db 3415 GAGATCTTCTGAAGGTGGTGGAGAGTGTGTGGGACACAGATATGAGAGTGGCAGC 3474
Qy 3690 TGGGGCAGCACCTATGACAGGCATTTGCTGGCTTAGACGTAAACCTTGGGCTCAAGATG 3749
Db 3475 TGGGGCAGCACCTATGACAGGCATTTGCTGGCTTAGACGTAAACCTTGGGCTCAAGATG 3534
Qy 3750 CCGCCACAGGAGACAGCCTGGAGAACCGGGGAAACAGCTGGGTGAGCCCGACAGACTGAC 3809
Db 3535 CCGCCACAGGAGACAGCCTGGAGAACCGGGGAAACAGCTGGGTGAGCCCGACAGACTGAC 3594
Qy 3810 CAGGCTCTGGGCCAGAGCCGTCGGCGGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3869
Db 3595 CAGGCTCTGGGCCAGAGCCGTCGGCGGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3654
Qy 3870 CTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGCCTCCGCGCAGCCCGCGGCTGTTC 3929
Db 3655 CTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGCCTCCGCGCAGCCCGCGGCTGTTC 3714
Qy 3930 GCCAGATGCTGCTGCTGCTCTTGTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 3989
Db 3715 GCCAGATGCTGCTGCTGCTCTTGTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 3774
Qy 3990 CCTCTTTCGGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTG 4049
Db 3775 CCTCTTTCGGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTG 3834
Qy 4050 TCTTCTTTCAGTAGGACCCCGAGGACCCCTGAGCGTGGCCGCTGCTCGAGCGCTG 4109
Db 3835 TCTTCTTTCAGTAGGACCCCGAGGACCCCTGAGCGTGGCCGCTGCTCGAGCGCTG 3894
Qy 4110 CTGACGAGGACGAGCTGGAGAGCCCGCAGTGCAGCATAGCTCCACAGGTTCTCGGCA 4169
Db 3895 CTGACGAGGACGAGCTGGAGAGCCCGCAGTGCAGCATAGCTCCACAGGTTCTCGGCA 3954

Qy 4170 CCAGAAGTTCTGTCTGTAAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGAACCCAGAGTCT 4229
Db 3955 CCAGAAGTTCTGTCTGTAAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGAACCCAGAGTCT 4014
Qy 4230 CCATCCCGCAGCTGCGCAGTGTAGCAGCCCGGTGGCCGCGCTGCTGCCCGACTGCCCG 4289
Db 4015 CCATCCCGCAGCTGCGCAGTGTAGCAGCCCGGTGGCCGCGCTGCTGCCCGACTGCCCG 4074
Qy 4290 GCTGAGCTGTGTGCTCCCTCCGCCCCAGCAGTGTAGCAGCCCGGTGGGGAGTGTTCAG 4349
Db 4075 GCTGAGCTGTGTGCTCCCTCCGCCCCAGCAGTGTAGCAGCCCGGTGGGGAGTGTTCAG 4134
Qy 4350 AACCTGACAGGCCGGAACCTGTCTGACTTCTGTGTTCAAGACTACCCGCGCTGTGTGGC 4409
Db 4135 AACCTGACAGGCCGGAACCTGTCTGACTTCTGTGTTCAAGACTACCCGCGCTGTGTGGC 4194
Qy 4410 CAGGCTGTAAGACTAAGAGTGGGTGAATGAGGTGAGGTACGAGGTCTTCTGCTGGG 4469
Db 4195 CAGGCTGTAAGACTAAGAGTGGGTGAATGAGGTGAGGTACGAGGTCTTCTGCTGGG 4254
Qy 4470 GGCAGAGACCCAGGCTGCGCTCGGGCCAGAGTTGGGCGCTCAGTGGAGAGTGTGG 4529
Db 4255 GGCAGAGACCCAGGCTGCGCTCGGGCCAGAGTTGGGCGCTCAGTGGAGAGTGTGG 4314
Qy 4530 GCGCTGTGAGTCCCTGCTGGCGGGGCCCTCGAACCGTGTCTTGAAAAAACCCTCACAGCC 4589
Db 4315 GCGCTGTGAGTCCCTGCTGGCGGGGCCCTCGAACCGTGTCTTGAAAAAACCCTCACAGCC 4374
Qy 4590 TGGGCTCACAGCTGGATGTCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGG 4649
Db 4375 TGGGCTCACAGCTGGATGTCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGG 4434
Qy 4650 CACTCCATGGTGGCTTTGTCAAACCGAGCCAGCAACGCAATCTCCGCTGCTCACCTGCC 4709
Db 4435 CACTCCATGGTGGCTTTGTCAAACCGAGCCAGCAACGCAATCTCCGCTGCTCACCTGCC 4494
Qy 4710 CCAGGCCCGGCCCGCCACGACATCAGATCAACACACTCAACACCCCTTGAAACCTCAC 4769
Db 4495 CCAGGCCCGGCCCGCCACGACATCAGATCAACACACTCAACACCCCTTGAAACCTCAC 4554
Qy 4770 AAGAGACGCTGTCTGAGGCTGCACTGATGCGCTCTCGGTGGAGCTCTCGTCTCCATC 4829
Db 4555 AAGAGACGCTGTCTGAGGCTGCACTGATGCGCTCTCGGTGGAGCTCTCGTCTCCATC 4614
Qy 4830 TGTGTGCTTTGGCCATGCTCTTGTCCGCGCAGCTTCACTTGTCTCATTTGAGAG 4889
Db 4615 TGTGTGCTTTGGCCATGCTCTTGTCCGCGCAGCTTCACTTGTCTCATTTGAGAG 4674
Qy 4890 CGAGTCAACCCGAGCCAAAGCACCTGCAGCTCATGGGGGCTGTCTCCCGACCTCTACTGG 4949
Db 4675 CGAGTCAACCCGAGCCAAAGCACCTGCAGCTCATGGGGGCTGTCTCCCGACCTCTACTGG 4734
Qy 4950 CTTCGCAACTTTCTCTGGGACATGTAACTAATTGGTGGCAGCATGATCGTGGTGTTC 5009
Db 4735 CTTCGCAACTTTCTCTGGGACATGTAACTAATTGGTGGCAGCATGATCGTGGTGTTC 4794
Qy 5010 ATCTTTCTGGCTTTCAGCAGAGGCGCATATGTGGCGCTGCAACCTGCTCTCTG 5069
Db 4795 ATCTTTCTGGCTTTCAGCAGAGGCGCATATGTGGCGCTGCAACCTGCTCTCTG 4854
Qy 5070 CTGTGTCTACTACTGTATGCTGCTGATCACCGCTCATGTACCCAGCCCTCTTCTTC 5129
Db 4855 CTGTGTCTACTACTGTATGCTGCTGATCACCGCTCATGTACCCAGCCCTCTTCTTC 4914
Qy 5130 TTCTCCGTGGCCAGCAGAGCTATGTGGTGTCTACCTGCAATAAACCCTTTATTGGGATC 5189
Db 4915 TTCTCCGTGGCCAGCAGAGCTATGTGGTGTCTACCTGCAATAAACCCTTTATTGGGATC 4974
Qy 5190 AATGGAGCAGTGGCCACTTTGTGCTTGTGAGCTTCTCTGATCAGAGCTGCAGAGGTG 5249
Db 4975 AATGGAGCAGTGGCCACTTTGTGCTTGTGAGCTTCTCTGATCAGAGCTGCAGAGGTG 5034

QY 5250 AGCGGATCTTGAACAGGCTTCTCTTATCTTCCCCACATCTTGTGCTTGGCCGCGGGCTC 5309
DB 5035 AGCGGATCTTGAACAGGCTTCTCTTATCTTCCCCACATCTTGTGCTTGGCCGCGGGCTC 5094
QY 5310 ATTGACATGTGCGGAACACAGGCCATGGCTGTATGCTTTCAGCGCTTGGAGACAGGCAG 5369
DB 5095 ATTGACATGTGCGGAACACAGGCCATGGCTGTATGCTTTCAGCGCTTGGAGACAGGCAG 5154
QY 5370 TTCCAGTCACTTCCCTGCGCTGGAGGTGGTGGCAAGAACCTTTCAGCATGGTGATACAG 5429
DB 5155 TTCCAGTCACTTCCCTGCGCTGGAGGTGGTGGCAAGAACCTTTCAGCATGGTGATACAG 5214
QY 5430 GGGCCCTCTTCTTCTTCACTACTCTGCTGAGCAGCAGCAAGCAACTCTTCCGCCACAG 5489
DB 5215 GGGCCCTCTTCTTCTTCACTACTCTGCTGAGCAGCAGCAAGCAACTCTTCCGCCACAG 5274
QY 5490 CCAGGGTGTAGGTCTCTGCGCACTCTTGGGAGAGGACAGGATGTAGCCCTGGAACGG 5549
DB 5275 CCAGGGTGTAGGTCTCTGCGCACTCTTGGGAGAGGACAGGATGTAGCCCTGGAACGG 5334
QY 5550 GAGCGGTGTCTCAAGAGGACCCAGGGGATGTGTGTGCTGGGATTTCCCTCTGCT 5609
DB 5335 GAGCGGTGTCTCAAGAGGACCCAGGGGATGTGTGTGCTGGGATTTCCCTCTGCT 5394
QY 5610 GTATACCGTGGGAGAGGATGTGTGTGCTGGGATTTCCCTCTGCT 5669
DB 5395 GTATACCGTGGGAGAGGATGTGTGTGCTGGGATTTCCCTCTGCT 5454
QY 5670 GAGTGTGTTTGGGCTGTGCTGTGTAATGGAGAGGGAAGACGTCACGTTTTCGATGGTG 5729
DB 5455 GAGTGTGTTTGGGCTGTGCTGTGTAATGGAGAGGGAAGACGTCACGTTTTCGATGGTG 5514
QY 5730 ACGGGGACACATTTGGCCAGAGGCGAGGCTGTGTGCTGGGACACAGCGTGGCCCGG 5789
DB 5515 ACGGGGACACATTTGGCCAGAGGCGAGGCTGTGTGCTGGGACACAGCGTGGCCCGG 5574
QY 5790 GAACCCAGTCTCGGACCTCAGATGGGATCTGCTCAATCCGATGCCATCTTTGAG 5849
DB 5575 GAACCCAGTCTCGGACCTCAGATGGGATCTGCTCAATCCGATGCCATCTTTGAG 5634
QY 5850 CTGCTGACGCGGCGGACGACCTTGGAGCTCTTTCGCGCGCTGCGCGGTGTCCGAGGCGC 5909
DB 5635 CTGCTGACGCGGCGGACGACCTTGGAGCTCTTTCGCGCGCTGCGCGGTGTCCGAGGCGC 5694
QY 5910 CAGGTTGCCAGACCGCTGCTCAGGCTGTGCGGCTGTGGGACTCTCATGTTACGACAGC 5969
DB 5695 CAGGTTGCCAGACCGCTGCTCAGGCTGTGCGGCTGTGGGACTCTCATGTTACGACAGC 5754
QY 5970 CGGCTGACGACACCTACAGCGAGGGAACAAACGCAAGCTGGCGAGCGGCTTGGCGCTG 6029
DB 5755 CGGCTGACGACACCTACAGCGAGGGAACAAACGCAAGCTGGCGAGCGGCTTGGCGCTG 5814
QY 6030 GTTGGGACCCAGCGGTGTGTTTCTGGACGAGCCGACACAGGCGATGGACCCCGAGCGG 6089
DB 5815 GTTGGGACCCAGCGGTGTGTTTCTGGACGAGCCGACACAGGCGATGGACCCCGAGCGG 5874
QY 6090 CGGCGCTTCTTGGAAACAGCCTTTTGGCGGTGTGGGAGGCGGCTTTCAGTGTGCTC 6149
DB 5875 CGGCGCTTCTTGGAAACAGCCTTTTGGCGGTGTGGGAGGCGGCTTTCAGTGTGCTC 5934
QY 6150 ACCTCCCATAGCATGGAGGATGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTTGAAT 6209
DB 5935 ACCTCCCATAGCATGGAGGATGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTTGAAT 5994
QY 6210 GGGCGGTTCGCTGCTGGGACCGGCAACATCTCAAGGGCAGATTCGCGGCGGTTCAC 6269
DB 5995 GGGCGGTTCGCTGCTGGGACCGGCAACATCTCAAGGGCAGATTCGCGGCGGTTCAC 6054
QY 6270 AACTGACCTTGGGGTTCGCGGCGGCAAGTCCAGCGGACGCGCTTTCGTGGCGGC 6329
DB 6055 AACTGACCTTGGGGTTCGCGGCGGCAAGTCCAGCGGACGCGCTTTCGTGGCGGC 6114
QY 6330 GAGTTCCTCGGCTGCGAGCTGCGGAGGACATGGAGGCGGCTTCCAGCTGCGG 6389

DB 6115 GAGTTCCTCGGCTCGAGCTGCGGAGGACATGGAGGTCGCTTCCAGCTGCGG 6174
QY 6390 CCGGAGGGGCGTGCCTTGGCGCGCTTGGAGAGCTGCGGCTGCAACGCGCGAGAG 6449
DB 6175 CCGGAGGGGCGTGCCTTGGCGCGCTTGGAGAGCTGCGGCTGCAACGCGCGAGAG 6234
QY 6450 CACGGCGTGGAGGACTTTTCCGTGAGCCAGACGATCTGGAGAGGTATTTCTGTACTTC 6509
DB 6235 CACGGCGTGGAGGACTTTTCCGTGAGCCAGACGATCTGGAGAGGTATTTCTGTACTTC 6294
QY 6510 TCCAGAGCACAGGGGAGGACGAGGACACCGAAGAGACAGAGGAGGAGTGGAGTG 6569
DB 6295 TCCAGAGCACAGGGGAGGACGAGGACACCGAAGAGACAGAGGAGGAGTGGAGTG 6354
QY 6570 GACCCGCGCGGCTGCGAGCACCCCAACCGCGTCAGCGTTCCTCGATGACCTTAGC 6629
DB 6355 GACCCGCGCGGCTGCGAGCACCCCAACCGCGTCAGCGTTCCTCGATGACCTTAGC 6414
QY 6630 ACTGCCGAGACTGTGCTC 6647
DB 6415 ACTGCCGAGACTGTGCTC 6432

RESULT 8
ABX72257
ID ABX72257 standard; cDNA; 6327 BP.
XX
AC ABX72257;
XX
DT 03-JUN-2003 (first entry)
XX
DE Human NOVX polynucleotide #88.
XX
KW Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
XX
XX WO200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0283512P.
PR 12-APR-2001; 2001US-0283710P.
PR 13-APR-2001; 2001US-0284234P.
PR 17-APR-2001; 2001US-0285325P.
PR 19-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.

PR	19-JUN-2001; 2001US-0299237P.	Db	241	TGCTTTCCGAGCTGACACCGGGGAGAGAGCCGGGGCCCTGAGCAACTTCAACGACTCC	300
PR	19-JUN-2001; 2001US-0299276P.	Qy	510	CTGGTCTCCCGGCTGTAGCCGATGCCCGCACTGTGTGGAGGGGGCAGTGCACACAGG	569
PR	12-SEP-2001; 2001US-0318750P.	Db	301	CTGGTCTCCCGGCTGTAGCCGATGCCCGCACTGTGTGGAGGGGGCAGTGCACACAGG	360
PR	25-SEP-2001; 2001US-0324800P.	Qy	570	ACGCTGGCTGCTAGGGAAGCTGATGCCACGCTGAGGGCTGCACGAGCAACGGCCGAG	629
PR	27-SEP-2001; 2001US-0325684P.	Db	361	ACGCTGGCTGCTAGGGAAGCTGATGCCACGCTGAGGGCTGCACGAGCAACGGCCGAG	420
PR	17-OCT-2001; 2001US-0330143P.	Qy	630	CCTCAACCAACCAAGAGCTCTCACTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	689
PR	14-NOV-2001; 2001US-0332131P.	Db	421	CCTCAACCAACCAAGAGCTCTCACTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	480
PR	14-NOV-2001; 2001US-0332240P.	Qy	690	ACGTCACTGTGCGCAGCGAATCCCTGGGGTTGGCACTGGGGCCCAAGCCGAGGAGCCCTTG	749
PR	21-NOV-2001; 2001US-0332115P.	Db	481	ACGTCACTGTGCGCAGCGAATCCCTGGGGTTGGCACTGGGGCCCAAGCCGAGGAGCCCTTG	540
PR	04-DEC-2001; 2001US-0337211P.	Qy	750	CACAGCTTGTGGAGGCGCTGAGGACCTGCGCCACGAGAGCTCTCTGGCGCTGGGAGCCCTG	809
PR	03-JAN-2002; 2002US-0345783P.	Db	541	CACAGCTTGTGGAGGCGCTGAGGACCTGCGCCACGAGAGCTCTCTGGCGCTGGGAGCCCTG	600
PR	16-JAN-2002; 2002US-0350251P.	Qy	810	GTGAGGCTTCCGGGCACTGCTCAGAGACCCCGAGGACCCAGCGGACCCAGCGGCTTGGATG	869
PR	02-APR-2002; 2002US-00114270.	Db	601	GTGAGGCTTCCGGGCACTGCTCAGAGACCCCGAGGACCCAGCGGCTTGGATG	660
XX	(CURA-) CURAGEN CORP.	Qy	870	TCAGAGGCCCTCTGCAGTGTGAGGAGCCTAGCAGCAGCAGTGGGGCCCTCCCTCAACTGG	929
XX	Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;	Db	661	TCAGAGGCCCTCTGCAGTGTGAGGAGCCTAGCAGCAGCAGTGGGGCCCTCCCTCAACTGG	720
PI	Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerkhusen BD;	Qy	930	TACAGGCTAGTACCTGATGAGAGCTGTGGGAGGAGCCAGAAATCCGCTTGCAGAC	989
PI	Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;	Db	721	TACAGGCTAGTACCTGATGAGAGCTGTGGGAGGAGCCAGAAATCCGCTTGCAGAC	780
PI	Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;	Qy	990	AGCAGCTGAGCCCGGCTGCTCGAGCTGATGGAGCCCTGGAGCCCTGGACAGCAGCCGCTGCC	1049
PI	Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;	Db	781	AGCAGCTGAGCCCGGCTGCTCGAGCTGATGGAGCCCTGGAGCCCTGGACAGCAGCCGCTGCC	840
PI	Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;	Qy	1050	CGCTGCTCTGGAGACGCTGAAAGCTCTGATCTCTGGGAAAGCTACTCTTTGACACAGAT	1109
PI	Ellerman K;	Db	841	CGCTGCTCTGGAGACGCTGAAAGCTCTGATCTCTGGGAAAGCTACTCTTTGACACAGAT	900
XX	WPI; 2003-046858/04.	Qy	1110	ACACCTTTTACCCGGAGCTCATGGCCAGGTGAACCGGACCTTCGAGAGAGCTCAACCTG	1169
XX	P-PSDB; ABUS4629.	Db	901	ACACCTTTTACCCGGAGCTCATGGCCAGGTGAACCGGACCTTCGAGAGAGCTCAACCTG	960
CC	New isolated NOVX polypeptide useful for treating atherosclerosis,	Qy	1170	CTGAGGATGTCGGGAGGTGTGGAGATGTGGGACCCCGGATCTTCACTTCAATGAC	1229
CC	metabolic disorders, diabetes, obesity, infectious disease, anorexia,	Db	961	CTGAGGATGTCGGGAGGTGTGGAGATGTGGGACCCCGGATCTTCACTTCAATGAC	1020
CC	neurodegenerative disorders, Alzheimer's disease and cancer.	Qy	1230	GACAGTTTCAATGTGGCCATGCTGCAGCGGCTCTCTGCAGATGAGGAGGAGGAGG	1289
PS	Claim 17; Page 280-281; 666pp; English.	Db	1021	GACAGTTTCAATGTGGCCATGCTGCAGCGGCTCTCTGCAGATGAGGAGGAGGAGG	1080
XX	The invention relates to human polypeptides, termed NOVX, and the	Qy	1290	CAGCCAGAGCTGAGGCGGGACCAATGAGGCGCTCGCATCTTTTCTGAGACCTCGG	1349
CC	polynucleotides encoding them. The polypeptides and polynucleotides	Db	1081	CAGCCAGAGCTGAGGCGGGACCAATGAGGCGCTCGCATCTTTTCTGAGACCTCGG	1140
CC	are useful for diagnosing disease, and screening for potential therapeutic	Qy	1350	AGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCGACCTGGTGGGCGAGCTG	1409
CC	agents. The sequences are useful for treating metabolic disorders,	Db	1141	AGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCGACCTGGTGGGCGAGCTG	1200
CC	cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic	Qy	1410	GGCGAGTGCAGGAGTGCCTGCTTGGACAAGCTGGAGGGCGGACCTCAGAGGACGCC	1469
CC	stenosis, atrial septal defect (ASD), atriocentric canal defect,	Db	1201	GGCGAGTGCAGGAGTGCCTGCTTGGACAAGCTGGAGGGCGGACCTCAGAGGACGCC	1260
CC	ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular	Qy	1470	CTGGTGTGCGGGCCCTGCAACTGCTCGCGAAACATCGATTCTGGGCGGGCTGTCTTC	1529
CC	septal defect (VSD), valve diseases, tuberosclerosis, scleroderma,	Db	1261	CTGGTGTGCGGGCCCTGCAACTGCTCGCGAAACATCGATTCTGGGCGGGCTGTCTTC	1320
CC	atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative	Qy	1530	TTGGGACCTGAGGAGCTCTTTCAGACCCCAAGAGACCCCAACCCAGACTCTGGGCGGGC	1589
CC	disorders, Alzheimer's disease, Parkinson's disease, immune disorders,	Db	1321	TTGGGACCTGAGGAGCTCTTTCAGACCCCAAGAGACCCCAACCCAGACTCTGGGCGGGC	1380
CC	haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease				
CC	and cancer. Sequences ABX72170-ABX72275 represent human NOVX				
CC	polynucleotides of the invention				
SQ	Sequence 6327 BP; 1052 A; 2054 C; 2059 G; 1162 T; 0 U; 0 Other;				
	Query Match 88.7%; Score 6036.8; DB 8; Length 6327;				
	Best Local Similarity 95.9%; Pred. No. 0;				
	Mismatches 6315; Conservative 0; Mismatches 12; Indels 261; Gaps 1;				
Qy	210 ATGGCTTCTGGACACAGCTGATGCTGTGCTGCTGGAAGAATTTTCAATGATCGCGGAGA	269			
Db	1 ATGGCTTCTGGACACAGCTGATGCTGTGCTGCTGGAAGAATTTTCAATGATCGCGGAGA	60			
Qy	270 CAGCGGTCCAGCTCCTGGTGGAAATGCTGTGGCCCTCTCTTCTTCTTCTTCTTCTTCTG	329			
Db	61 CAGCGGTCCAGCTCCTGGTGGAAATGCTGTGGCCCTCTCTTCTTCTTCTTCTTCTTCTG	120			
Qy	330 GCTGTTGCCACTCCACCGGCTTGGAGCAGCATGATGCCACTTCCCAACAGGCA	389			
Db	121 GCTGTTGCCACTCCACCGGCTTGGAGCAGCATGATGCCACTTCCCAACAGGCA	180			
Qy	390 CTGCATCGGGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	449			
Db	181 CTGCATCGGGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240			
Qy	450 TGCTTTCCGAGCTGACACCGGGGAGGAGCCCGGCGCTTGGAGCAACTTCAACGACTCC	509			

QY 5970 CGGCTGACGACCTACAGCGAGGGAACAAACGCAAGCTGGCGCCCTGGCGCTG 6029
DB 5500 CGGCTGACGACCTACAGCGAGGGAACAAACGCAAGCTGGCGCCCTGGCGCTG 5559
QY 6030 GTTGGGACACGACCGGTGTTTCTGGAGAGCCGACCAACAGGCATGGACCCAGCGG 6089
DB 5560 GTTGGGACACGACCGGTGTTTCTGGAGAGCCGACCAACAGGCATGGACCCAGCGG 5619
QY 6090 CGGCGCTTCTTTGGAAACAGCCTTTTGGCCGCTGGTGGGAGGCGCTTCAAGTATGCTC 6149
DB 5620 CGGCGCTTCTTTGGAAACAGCCTTTTGGCCGCTGGTGGGAGGCGCTTCAAGTATGCTC 5679
QY 6150 ACCTCCATAGATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGGTGAAT 6209
DB 5680 ACCTCCATAGATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGGTGAAT 5739
QY 6210 GGGCGGTTCGCTGCTGGGAGCGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCA 6269
DB 5740 GGGCGGTTCGCTGCTGGGAGCGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCA 5799
QY 6270 ACATGACCTCGCGGTGCGCGCGCAAGTCCAGCGCGAGCGGCTTCTGTTGGCGCC 6329
DB 5800 ACATGACCTCGCGGTGCGCGCGCAAGTCCAGCGCGAGCGGCTTCTGTTGGCGCC 5859
QY 6330 GAGTTCCTCGGTGCGAGCTGCGGAGCGACATGGAGCGCGCTTCCAGCTGCGG 6389
DB 5860 GAGTTCCTCGGTGCGAGCTGCGGAGCGACATGGAGCGCGCTTCCAGCTGCGG 5919
QY 6390 CCGGAGGGGCGCTGCGCGCTGGCGCGCTTTTGGAGAGCTGGCGGTCAGCGCGCAGAG 6449
DB 5920 CCGGAGGGGCGCTGCGCGCTGGCGCGCTTTTGGAGAGCTGGCGGTCAGCGCGCAGAG 5979
QY 6450 CAGCGCTGAGGACTTTTCCGTGAGCCAGACAGATGCTGAGGAGTATTTCTGTACTTC 6509
DB 5980 CAGCGCTGAGGACTTTTCCGTGAGCCAGACAGATGCTGAGGAGTATTTCTGTACTTC 6039
QY 6510 TCCAGGACACAGGGAGGAGGAGGACACCAAGAGCAGAGGAGGAGGAGGAGTG 6569
DB 6040 TCCAGGACACAGGGAGGAGGAGGACACCAAGAGCAGAGGAGGAGGAGGAGTG 6099
QY 6570 GACCCCGCGCGAGCTGACGACCCCAAGAGGCTGAGGAGTTCCTCGATGACCCCTAGC 6629
DB 6100 GACCCCGCGCGAGCTGACGACCCCAAGAGGCTGAGGAGTTCCTCGATGACCCCTAGC 6159
QY 6630 ACTGCCGAGACTGTGCTCTGAGCTTCCCTCCCTGCGGGGCGCGGGAGGCCCTGGGAA 6689
DB 6160 ACTGCCGAGACTGTGCTCTGAGCTTCCCTCCCTGCGGGGCGCGGGAGGCCCTGGGAA 6219
QY 6690 TGGCAAGGGCAAGGTAGTGCCTTAGAGCCCTGGAAGCTGAGGAGGAGGAGGAGTG 6749
DB 6220 TGGCAAGGGCAAGGTAGTGCCTTAGAGCCCTGGAAGCTGAGGAGGAGGAGGAGTG 6279
QY 6750 CCCTGGAGAAATAAAGAGAGGCTGGAGAGAGCCCTGCTGGTGA 6797
DB 6280 CCCTGGAGAAATAAAGAGAGGCTGGAGAGAGCCCTGCTGGTGA 6327

RESULT 9

ABX95283
ID ABX95283 standard; cdna; 6696 BP.
XX AC ABX95283;
XX
XX 18-JUN-2003 (first entry)
DE cdna encoding human ABCA7 splice variant #1.
XX Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic; gene; ss.

XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 673..6696
FT /*tag= a
FT /partial
FT /product= "ABCA7 splice variant #1"
FT /note= "this sequence lacks a stop codon. The ORF given
FT as SEQ ID No:2 is specifically claimed in Claim 4"
XX
PN WO2003010315-A1.
XX
XX 06-FEB-2003.
XX 24-JUL-2002; 2002WO-JP007487.
XX
XX 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
XX Ueda K, Nakagawa S, Nagase T;
PI
XX WPI: 2003-239444/23.
DR P-PSDB; ABU08464.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 4; Page 77-92; 183pp; Japanese.
XX
XX The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence encodes human ABCA7 splice variant #1
XX
SQ Sequence 6696 BP; 1103 A; 2172 C; 2153 G; 1268 T; 0 U; 0 Other;
Query Match 87.2%; Score 5933.2; DB 8; Length 6696;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5938; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 702 CGCACCGAATCCCTGGGGTTGGCACTGGGGCCCAAGCCAGAGCCCTTGACAGCTTTG 761
DB 751 CCCAGGATCCCTGGGGTTGGCACTGGGGCCCAAGCCAGAGCCCTTGACAGCTTTG 810
QY 762 GAGGCGCTGAGGACTTGGCCCAAGGAGCTTCTGGCGCTGCGAGCTTGTGGAGCTTCG 821
DB 811 GAGGCGCTGAGGACTTGGCCCAAGGAGCTTCTGGCGCTGCGAGCTTGTGGAGCTTCG 870
QY 822 GCATCTGTCAGAGACCCCGAGGAGCCAGGCGCCCTGGAGTTGCTGTCAGAGGCCCTC 881
DB 871 GCATCTGTCAGAGACCCCGAGGAGCCAGGCGCCCTGGAGTTGCTGTCAGAGGCCCTC 930
QY 882 TGCAGTGTGAGGGGAGCTAGCAGCAGACAGTGGGGCCCTCCCTCAACTGGTACGAGGCTAGT 941
DB 931 TGCAGTGTGAGGGGAGCTAGCAGCAGACAGTGGGGCCCTCCCTCAACTGGTACGAGGCTAGT 990
QY 942 GACCTGATGAGCTGTGTGGGGCAGGAGCCAGAAATCCGCCCTTGCACAGCAGCCTGAGC 1001
DB 991 GACCTGATGAGCTGTGTGGGGCAGGAGCCAGAAATCCGCCCTTGCACAGCAGCCTGAGC 1050
QY 1002 CCGGCTGCTCGGAGCTGATTGGAGCCCTGGAGCAGCCACCGCTGTCCCGCCCTGCTCGG 1061
DB 1051 CCGGCTGCTCGGAGCTGATTGGAGCCCTGGAGCAGCCACCGCTGTCCCGCCCTGCTCGG 1110
QY 1062 AGACGCTGAAGCCTCTGATTCCTCGGGAAGTACTCTTTGACACAGATACACCTTTTACC 1121

QY 3282 CTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCCGCCTGCCCTG 3341
DB CTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCCGCCTGCCCTG 3390
QY 3342 ACCACCAATGAGAAAGGCTGACACTGACATGAGAGGCGAGTGTGGACACACGAGGCAAGAAAG 3401
DB ACCACCAATGAGAAAGGCTGACACTGACATGAGAGGCGAGTGTGGACACACGAGGCAAGAAAG 3450
QY 3402 AAGNATGGCAGCCAGGCGAGGCTGAGAGTGGCACTCTCAGCTGCTGGCCCTGGTACAGCAC 3461
DB AAGNATGGCAGCCAGGCGAGGCTGAGAGTGGCACTCTCAGCTGCTGGCCCTGGTACAGCAC 3510
QY 3462 TGGGTGCCCGGGGCACGGCTGGTGGAGAGCTGCCACACAGAGCTGGTGTGCTGCC 3521
DB TGGGTGCCCGGGGCACGGCTGGTGGAGAGCTGCCACACAGAGCTGGTGTGCTGCC 3570
QY 3522 TACACGGGTCCCCATGACGCGAGCTTCGCCACACTCTTCGAGAGCTAGACACGCGGCTG 3581
DB TACACGGGTCCCCATGACGCGAGCTTCGCCACACTCTTCGAGAGCTAGACACGCGGCTG 3630
QY 3582 GCGAGCTGAGGCTCACTGGCTACGGGATCTCGACACGAGCTCGAGGAGATCTTCCTG 3641
DB GCGAGCTGAGGCTCACTGGCTACGGGATCTCGACACGAGCTCGAGGAGATCTTCCTG 3690
QY 3642 AAGTGTGGAGGAGTGTCTGGCGACACAGATATGAGGATGGCAGCTGCGGGGACGAC 3701
DB AAGTGTGGAGGAGTGTCTGGCGACACAGATATGAGGATGGCAGCTGCGGGGACGAC 3750
QY 3702 CTATGACAGGCAATGCTGGCTAGACGTAAACCTCGGGCTCAAGATGCGGCCACACAGGAG 3761
DB CTATGACAGGCAATGCTGGCTAGACGTAAACCTCGGGCTCAAGATGCGGCCACACAGGAG 3810
QY 3762 ACAGCGCTGAGAAACGGGGAAACAGCTGGGTGAGCCCCAGAGACTGACAGGGCTCTGGG 3821
DB ACAGCGCTGAGAAACGGGGAAACAGCTGGGTGAGCCCCAGAGACTGACAGGGCTCTGGG 3870
QY 3822 CCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCCCTG 3881
DB CCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCCCTG 3930
QY 3882 CTCTCAAGGCTTTCTGCTTGCCCGCGCAGCGCGCGGCTGTTCGCGCAGATCGTG 3941
DB CTCTCAAGGCTTTCTGCTTGCCCGCGCAGCGCGCGGCTGTTCGCGCAGATCGTG 3990
QY 3942 CTGCTGCCCTCTTTGTGGGCTGGGCCCTCGTGTTCAGCCTCAFGTGGCTCTCTTCGGG 4001
DB CTGCTGCCCTCTTTGTGGGCTGGGCCCTCGTGTTCAGCCTCAFGTGGCTCTCTTCGGG 4050
QY 4002 CACTACCGGCTCTGCGGCTCAGTCCACCATGTACGGTGTCTCAGGTGTCTCTTCAGT 4061
DB CACTACCGGCTCTGCGGCTCAGTCCACCATGTACGGTGTCTCAGGTGTCTCTTCAGT 4110
QY 4062 GAGGACGCCCCAGGGACCTTGACGTGCCCGGCTGCTCAGAGCGCTGTCTGACAGGAGCA 4121
DB GAGGACGCCCCAGGGACCTTGACGTGCCCGGCTGCTCAGAGCGCTGTCTGACAGGAGCA 4170
QY 4122 GGACTGAGAGAGCCCCAGTGCAGCATAGTTCOCACAGGTTTCGGCACACAGAGTTTCCT 4181
DB GGACTGAGAGAGCCCCAGTGCAGCATAGTTCOCACAGGTTTCGGCACACAGAGTTTCCT 4230
QY 4182 GCTGAAGTGGCCAGGTCCTTGGCCAGTGGCACTGGACCCAGAGTGTCTCCATCCCCAGCC 4241
DB GCTGAAGTGGCCAGGTCCTTGGCCAGTGGCACTGGACCCAGAGTGTCTCCATCCCCAGCC 4290
QY 4242 TGCAGTGTAGCCAGCCCGGTGCCCCGGCTGCTGCCCGACTGCCCGGCTGACAGTGTGT 4301
DB TGCAGTGTAGCCAGCCCGGTGCCCCGGCTGCTGCCCGACTGCCCGGCTGACAGTGTGT 4350
QY 4302 GGTCCCCCTCCGCCACAGGAGTGAACCGGTCTGTGGGAAAGTGTTCAGAAACCTGACAGGC 4361
DB GGTCCCCCTCCGCCACAGGAGTGAACCGGTCTGTGGGAAAGTGTTCAGAAACCTGACAGGC 4410

QY 4362 CGGAACCTGTCTGACTCTCTCTGTTCAAGACCTACCCGCGCTGGTGGCCAGGCGCTGAAG 4421
DB CGGAACCTGTCTGACTCTCTCTGTTCAAGACCTACCCGCGCTGGTGGCCAGGCGCTGAAG 4470
QY 4422 ACTAAGAAGTGGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCA 4481
DB ACTAAGAAGTGGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCCA 4530
QY 4482 GGCTGCCCTCGGGCCCAAGAGTTGGGCCCTCAGTGGAGAGTGTGGGGCTGCTGAGT 4541
DB GGCTGCCCTCGGGCCCAAGAGTTGGGCCCTCAGTGGAGAGTGTGGGGCTGCTGAGT 4590
QY 4542 CCCTGCTGCGGGGGCCCTCGACCGCTGTCTCAAAAACCTCAGAGCTGGGCTCAAGC 4601
DB CCCTGCTGCGGGGGCCCTCGACCGCTGTCTCAAAAACCTCAGAGCTGGGCTCAAGC 4650
QY 4591 CCCTGCTGCGGGGGCCCTCGACCGCTGTCTGAAAAAACTCAGAGCTGGGCTCAAGC 4661
DB CCCTGCTGCGGGGGCCCTCGACCGCTGTCTGAAAAAACTCAGAGCTGGGCTCAAGC 4710
QY 4662 GCCTTTGTCAACCGAGCGACGCAACCTCCGCTGCTCACCTGCCCGCCAGGCCCGGCC 4721
DB GCCTTTGTCAACCGAGCGACGCAACCTCCGCTGCTCACCTGCCCGCCAGGCCCGGCC 4770
QY 4722 CGCACGCGCCACAGCATCACACACTCAACACCCCTTGAACTCACCAGAGGACAGCTG 4781
DB CGCACGCGCCACAGCATCACACACTCAACACCCCTTGAACTCACCAGAGGACAGCTG 4830
QY 4782 TCTGAGGCTGCACTGATGAGCTCCTCGTGGAGCGTCTCGTCTCCATCTGTGTGGTCTTT 4841
DB TCTGAGGCTGCACTGATGAGCTCCTCGTGGAGCGTCTCGTCTCCATCTGTGTGGTCTTT 4890
QY 4842 GCATGTCCTTTGTCGGGCGAGCTTCACTCTGTGCTCACTTCAAGAGGAGGACGAGTCA 4901
DB GCATGTCCTTTGTCGGGCGAGCTTCACTCTGTGCTCACTTCAAGAGGAGGACGAGTCA 4950
QY 4902 GCCAAGCAGCTGAGCTCATGCGGGGCGCTGTCCGCCACCTCTACCTGCGCTTGGCACTTT 4961
DB GCCAAGCAGCTGAGCTCATGCGGGGCGCTGTCCGCCACCTCTACCTGCGCTTGGCACTTT 5010
QY 4962 CTCTGGGACATGTGTAACTACTTGTGGTGGCAGCATGATCGTGTGTCTCATCTTTCTGGCC 5021
DB CTCTGGGACATGTGTAACTACTTGTGGTGGCAGCATGATCGTGTGTCTCATCTTTCTGGCC 5070
QY 5022 TTCACGAGAGGCGCATATGCGGCCCTGCCAACCTGCTGCTCTCTGCTGTGCTACTA 5081
DB TTCACGAGAGGCGCATATGCGGCCCTGCCAACCTGCTGCTCTCTGCTGTGCTACTA 5130
QY 5082 CTGTATGGCTGGTTCGATCACACGCTCATGTACCCAGCCTCTCTCTCTCTCCGTGCC 5141
DB CTGTATGGCTGGTTCGATCACACGCTCATGTACCCAGCCTCTCTCTCTCTCCGTGCC 5190
QY 5142 AGCACAGCCTATGTGTGTCTCACTGCATAAACTCTTTATTGGCATCAATGGAAGCATG 5201
DB AGCACAGCCTATGTGTGTCTCACTGCATAAACTCTCTTTATTGGCATCAATGGAAGCATG 5250
QY 5202 GCCACCTTTGTGCTTGAAGCTTCTCTGATCAGAACTGAGAGGTGAGCCGAGCTTG 5261
DB GCCACCTTTGTGCTTGAAGCTTCTCTGATCAGAACTGAGAGGTGAGCCGAGCTTG 5310
QY 5262 AAACAGCTCTCTTATCTTCCCGCTCTGCTGGGCGGGGCTCATTTGACATGGTG 5321
DB AAACAGCTCTCTTATCTTCCCGCTCTGCTGGGCGGGGCTCATTTGACATGGTG 5370
QY 5322 CGGAACGAGGCCATGCTGCTTTTGTGCGCTTGGGAGACAGGAGTTCAGTCAACC 5381
DB CGGAACGAGGCCATGCTGCTGCTTTTGTGCGCTTGGGAGACAGGAGTTCAGTCAACC 5430
QY 5382 CTGCGCTGGGAGGTGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTTC 5441
DB CTGCGCTGGGAGGTGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTTC 5490
QY 5442 CTTCTCTTCACACTACTGCTGCGAGCCAGAAAGCAACTCTCTGCCACAGCCAGGGTGAGG 5501

Db	5491	CTTCTCTTTCACACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAGCCACCGGTTAGG	5550
Qy	5502	TCTCTGCCACTCTCTGGGAGAGGAGCAGAGGATAGCCCGTGAAACGGGAGCGGTTGGTC	5561
Db	5551	TCTCTGCCACTCTCTGGGAGAGGAGCAGAGGATAGCCCGTGAAACGGGAGCGGTTGGTC	5610
Qy	5562	CAAGGAGCACCCAGGGGGATGTGTTGGTGTGTAGGAACTTGTGACCAAGGTATACCGTGGG	5621
Db	5611	CAAGGAGCACCCAGGGGGATGTGTTGGTGTGTAGGAACTTGTGACCAAGGTATACCGTGGG	5670
Qy	5622	CAGAGGATGCCAGCTGTGTGACCGCTGTGCTCGGGGATCCCTCGTGTAGTGTGTTTGGG	5681
Db	5671	CAGAGGATGCCAGCTGTGTGACCGCTGTGCTCGGGGATCCCTCGTGTAGTGTGTTTGGG	5730
Qy	5682	CTGCTGGGTGTGAATGGAGCAGGGAAGACGTGCACAGTTCGCACTGTGTGACGGGGGACACA	5741
Db	5731	CTGCTGGGTGTGAATGGAGCAGGGAAGACGTGCACAGTTCGCACTGTGTGACGGGGGACACA	5790
Qy	5742	TTGCCACAGAGGGCGAGGCTGTGCTGTGCAGGCCACAGCGTGTGCCCGGGAACCCAGTGTCT	5801
Db	5791	TTGCCACAGAGGGCGAGGCTGTGCTGTGCAGGCCACAGCGTGTGCCCGGGAACCCAGTGTCT	5850
Qy	5802	GCGCACCTCAGCATGGGATGCTGCGCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5861
Db	5851	GCGCACCTCAGCATGGGATGCTGCGCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5910
Qy	5862	GCGAGCACCTGGAGCTGTGTGCGCGCTGTGCGGAGTGTCCCGAGGCCCAGGTTGCCCGAG	5921
Db	5911	GCGAGCACCTGGAGCTGTGTGCGCGCTGTGCGGAGTGTCCCGAGGCCCAGGTTGCCCGAG	5970
Qy	5922	ACCCTGGCTCAGGCTGTGGCGCTGTGGGACTCTCATGTGTACGAGACCGGCTGTGCAAGC	5981
Db	5971	ACCCTGGCTCAGGCTGTGGCGCTGTGGGACTCTCATGTGTACGAGACCGGCTGTGCAAGC	6030
Qy	5982	ACCTACAGCGGAGGGAACAAACGCAAGCTGGCGACGCGCCTGTGGCGCTGTGTTGGGACCCA	6041
Db	6031	ACCTACAGCGGAGGGAACAAACGCAAGCTGGCGACGCGCCTGTGGCGCTGTGTTGGGACCCA	6090
Qy	6042	GCGGTGTGTTCTGGACGAGCCGACACACAGGCATGACCCACAGCGCGCGCGCTTCCCT	6101
Db	6091	GCGGTGTGTTCTGGACGAGCCGACACACAGGCATGACCCACAGCGCGCGCGCTTCCCT	6150
Qy	6102	TGGAACAGCCTTTTGGCGTGTGTCGGGAGGGCGGTTTCAGTGATGTCTCACTCCCATAGC	6161
Db	6151	TGGAACAGCCTTTTGGCGTGTGTCGGGAGGGCGGTTTCAGTGATGTCTCACTCCCATAGC	6210
Qy	6162	ATGAGAGATGTGAAGCGCTGTGCTCGGCGCTTAGCCATCATGTGTGAATGGCGGTTCCGC	6221
Db	6211	ATGAGAGATGTGAAGCGCTGTGCTCGGCGCTTAGCCATCATGTGTGAATGGCGGTTCCGC	6270
Qy	6222	TGCTGTGGCAGCCGCAACATCTCAAGGCAGATTCGCGCGGCTCACACATGACCCCTG	6281
Db	6271	TGCTGTGGCAGCCGCAACATCTCAAGGCAGATTCGCGCGGCTCACACATGACCCCTG	6330
Qy	6282	CGGTTGCCCGCGCAAGGTCCAGCCGGCAGCGGCTTCTGTGGCGCGCGAGTTCCTTGGG	6341
Db	6331	CGGTTGCCCGCGCAAGGTCCAGCCGGCAGCGGCTTCTGTGGCGCGCGAGTTCCTTGGG	6390
Qy	6342	TCGAGCTGTGCGCGAGGCACATGAGAGCGCGCTGTGCGTTCCAGTGTGCGCGGAGGCGC	6401
Db	6391	TCGAGCTGTGCGCGAGGCACATGAGAGCGCGCTGTGCGTTCCAGTGTGCGCGGAGGCGC	6450
Qy	6402	TGCGCCTTGGCGCGCTTCTTGGAGAGCTGGCGGTGACCGCGCAGAGCACGCGGTGGAG	6461
Db	6451	TGCGCCTTGGCGCGCTTCTTGGAGAGCTGGCGGTGACCGCGCAGAGCACGCGGTGGAG	6510
Qy	6462	GACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTCTCCAAAGGACCA	6521
Db	6511	GACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTCTCCAAAGGACCA	6570
Qy	6522	GGGAAGGACGAGGACACCGAAGACGAGAGGCGAGGATGGGAGTTGGACCCCGGCGCA	6581

Db	6571	GGGAAGGACGAGGACACCGAAGAGCAGACGAGGAGTGGAGTGGAGACCCCGGCGCCA	6633
Qy	6582	GGCCTGAGCAGACCCCAACCGTCGAGCCAGTTCCTCGATGACCTAGCAGTGGCGAGACT	6641
Db	6631	GGCCTGAGCAGACCCCAACCGTCGAGCCAGTTCCTCGATGACCTAGCAGTGGCGAGACT	6690
Qy	6642	GTGCTC 6647	
Db	6691	GTGCTC 6696	
RESULT 10			
ABX95284			
ID	ABX95284	standard; cDNA; 6651 BP.	
XX	AC	ABX95284;	
XX	XX		
DT	18-JUN-2003	(first entry)	
DE	cDNA encoding human ABCA7 splice variant #2.		
XX	Human; ATP-binding cassette transporter protein A7; ABC transporter;		
KW	ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;		
KW	abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;		
KW	immunomodulator; immunosuppressive; antiinflammatory;		
KW	antiarteriosclerotic; gene; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	673..6651	
FT	FT	/*tag= a	
FT	FT	/partial	
FT	FT	/product= "ABCA7 splice variant #2"	
FT	FT	/note= "This sequence lacks a stop codon. The ORF given	
FT	FT	as SEQ ID No:5 is specifically claimed in Claim 4"	
XX	XX		
PN	WO2003010315-A1.		
XX	XX		
PD	06-FEB-2003.		
XX	XX		
PF	24-JUL-2002;	2002WO-JP007487.	
XX	XX		
PR	25-JUL-2001;	2001JP-00224176.	
PR	06-DEC-2001;	2001JP-00372530.	
XX	XX		
PA	(KYOW) KYOWA HAKKO KOGYO KK.		
PA	(KAZU-) KAZUSA DNA RES INST FOUND.		
XX	XX		
PI	Ueda K, Nakagawa S, Nagase T;		
XX	XX		
DR	WPI; 2003-239444/23.		
DR	P-FSDB; ABU08465.		
XX	XX		
PT	Novel ABC transporter protein, ABCA7 splicing variant, participating in		
PT	the immune system, applicable in diagnosis of and screening drugs for		
PT	e.g. autoimmune diseases, Sjogren's syndrome and inflammations.		
XX	XX		
PS	Claim 4; Page 116-130; 183pp; Japanese.		
XX	XX		
CC	The present invention relates to the isolation of human ATP-binding		
CC	cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the		
CC	polynucleotide sequences encoding them. The protein is applicable in the		
CC	diagnosis and screening of drugs for autoimmune diseases, Sjogren's		
CC	syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.		
CC	It may also be used in a method for screening ABCA-SSN inhibitors. The		
CC	present sequence encodes human ABCA7 splice variant #2		
XX	XX		
SQ	Sequence 6651 BP; 1096 A; 2155 C; 2138 G; 1262 T; 0 U; 0 Other;		
Query Match 85.7%; Score 5833.2; DB 8; Length 6651;			
Best Local Similarity 99.1%; Pred. No. 0;			
Matches 5893; Conservative 0; Mismatches 8; Indels 45; Gaps 1;			

QY	702	CGCAGGGAATCCCTGGGTTGGCACTGGGCGCAAGCCAGGAGCCCTTTGCACAGCTTGTG	761
Db	751	CCCGAGGAATCCCTGGGTTGGCACTGGGCGCAAGCCAGGAGCCCTTTGCACAGCTTGTG	810
QY	762	GAGGCGCTGAGGACCTGGGCCCAAGGAGCTCTCGCGCTGCGCAGCCCTGTGTGAGCTTCGG	821
Db	811	GAGGCGCTGAGGACCTGGGCCCAAGGAGCTCTCGCGCTGCGCAGCCCTGTGTGAGCTTCGG	870
QY	822	GCACTGCTGAGAGACCCCGAGGGAACGAGGCGCCCTTGAGATTGCTGACAGGCGCTC	881
Db	871	GCACTGCTGAGAGACCCCGAGGGAACGAGGCGCCCTTGAGATTGCTGACAGGCGCTC	930
QY	882	TGCAGTGTACAGGAGCCTAGCAGCACAGTGGGCGCCCTCCCTCAACTGCTACAGGCTAGT	941
Db	931	TGCAGTGTACAGGAGCCTAGCAGCACAGTGGGCGCCCTCCCTCAACTGCTACAGGCTAGT	990
QY	942	GACCTGATGAGTGTGGGGCAGGAGCCAGAAATCGCGCCCTGCCAGACAGCAGCCTGAGC	1001
Db	991	GACCTGATGAGTGTGGGGCAGGAGCCAGAAATCGCGCCCTGCCAGACAGCAGCCTGAGC	1050
QY	1002	CCGCGCTGCTCGAGCTGATTGAGGCTTGGAAGCCACCCGCTGTCCCGCCCTGCTCTGG	1061
Db	1051	CCGCGCTGCTCGAGCTGATTGAGGCTTGGAAGCCACCCGCTGTCCCGCCCTGCTCTGG	1110
QY	1062	AGAGCGCTGAAGCCTCTGATCCTCGGAGAGTACTCTTTGCCACAGATACACCTTTTACC	1121
Db	1111	AGAGCGCTGAAGCCTCTGATCCTCGGAGAGTACTCTTTGCCACAGATACACCTTTTACC	1170
QY	1122	CGGAAGCTCATGTGCCCAAGGTGAACCGGACCTTCGAGAGCTCACCCCTGCTGAGGGATGTC	1181
Db	1171	CGGAAGCTCATGTGCCCAAGGTGAACCGGACCTTCGAGAGCTCACCCCTGCTGAGGGATGTC	1230
QY	1182	CGGAGGTGTGGAGATGCTTGGACCCCGGATCTTCACTTCATGAACGACAGTTCCAAT	1241
Db	1231	CGGAGGTGTGGAGATGCTTGGACCCCGGATCTTCACTTCATGAACGACAGTTCCAAT	1290
QY	1242	GTGGCCATGCTGAGCGGCTCTCGAGATGCGAGATGACAGATGAGGAGAGGAGCCAGACCT	1301
Db	1291	GTGGCCATGCTGAGCGGCTCTCGAGATGCGAGATGACAGATGAGGAGAGGAGCCAGACCT	1350
QY	1302	GGAGGCGGGACACATGAGGCGCTCGGATCCTTTCTGGACCTGGGAGCGGTGGCTAC	1361
Db	1351	GGAGGCGGGACACATGAGGCGCTCGGATCCTTTCTGGACCTGGGAGCGGTGGCTAC	1410
QY	1362	AGCTGGCAGGACGCAACGCTGATGTGGGCACTGTGGTGGACGCTGGGCGGAGTGACG	1421
Db	1411	AGCTGGCAGGACGCAACGCTGATGTGGGCACTGTGGTGGACGCTGGGCGGAGTGACG	1470
QY	1422	GAGTGGCTGTCTTGGACAGCTGGAGGCGGCAACCTTCAGAGGAGCCCTGGTGTGCGCG	1481
Db	1471	GAGTGGCTGTCTTGGACAGCTGGAGGCGGCAACCTTCAGAGGAGCCCTGGTGTGCGCG	1530
QY	1482	GCCTTGCAACTGCTCGCGGAACATCGATTCTGGGCGCGGCTCGTCTTTCTGGGACCTGAG	1541
Db	1531	GCCTTGCAACTGCTCGCGGAACATCGATTCTGGGCGCGGCTCGTCTTTCTGGGACCTGAG	1590
QY	1542	GACTCTTCAGACCCCAAGAGCACCAACCCAGACCTGGGCGCCCGGCCACAGTGGCATC	1601
Db	1591	GACTCTTCAGACCCCAAGAGCACCAACCCAGACCTGGGCGCCCGGCCACAGTGGCATC	1650
QY	1602	AAATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1661
Db	1651	AAATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1710
QY	1662	GACCTTGGCCAGCGCGGAGCCCTGACCGACTGCGCTTACGCTGAGGCGCGCTTCGTG	1721
Db	1711	GACCTTGGCCAGCGCGGAGCCCTGACCGACTGCGCTTACGCTGAGGCGCGCTTCGTG	1770
QY	1722	TACCTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCTAGCGCGCGCAACCCCGCG	1781
Db	1771	TACCTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCTAGCGCGCGCGCAACCCCGCG	1830

QY	1782	GCCGGCTCTTACTGTCAGCAGATGCCCTATCCGTGTATGTGACGACGCTGTTCCTGCGT	1841
Db	1831	GCCGGCTCTTACTGTCAGCAGATGCCCTATCCGTGTATGTGACGACGCTGTTCCTGCGT	1890
QY	1842	GTGCTGAGCCGGTCCGTGCGCTCTTCTGACGCTGGCTGGATCTACTCCGTGACACTG	1901
Db	1891	GTGCTGAGCCGGTCCGTGCGCTCTTCTGACGCTGGCTGGATCTACTCCGTGACACTG	1950
QY	1902	ACAGTGAAGCCGTGTGTGCGGAGAGGAGACGCGCTGCGGACACCATGCGCGGCATG	1961
Db	1951	ACAGTGAAGCCGTGTGTGCGGAGAGGAGACGCGCTGCGGACACCATGCGCGGCATG	2010
QY	1962	GCGCTCAGCGCGCGCTGCTCTGGCTAGGCTGGTCTCAGCTGCTCGGGCCCTTCCTG	2021
Db	2011	GCGCTCAGCGCGCGCTGCTCTGGCTAGGCTGGTCTCAGCTGCTCGGGCCCTTCCTG	2070
QY	2022	CTCAGCGCCGCTGCTGGTCTCTGGTCTCAAGCTGGGAGCATCTCCCTACAGCCAC	2081
Db	2071	CTCAGCGCCGCTGCTGGTCTCTGGTCTCAAGCTGGGAGCATCTCCCTACAGCCAC	2130
QY	2082	CGGGGCTGGTCTTCTCTGTTCTTGGCAGCTTTCGGCTGGCCACGCTGACCCAGAGCTTC	2141
Db	2131	CGGGGCTGGTCTTCTCTGTTCTTGGCAGCTTTCGGCTGGCCACGCTGACCCAGAGCTTC	2190
QY	2142	CTGCTCAGCGCCTTCTTCTCCCGCGCAACCTGGCTGCGGCTGCGCGCGCTTCGGCTAC	2201
Db	2191	CTGCTCAGCGCCTTCTTCTCCCGCGCAACCTGGCTGCGGCTGCGCGCGCTTCGGCTAC	2250
QY	2202	TTCTCCCTCTACTGCTGCTTCTGCTGTGTGGCTTGGCGGACCGGCTGCCCGCGGT	2261
Db	2251	TTCTCCCTCTACTGCTGCTTCTGCTGTGTGGCTTGGCGGACCGGCTGCCCGCGGT	2310
QY	2262	GCGCGCTGGCGCGGAGCCTGCTGCTGCGCGCTGCGCTTGGCTTGGCTTGGCGGCTG	2321
Db	2311	GCGCGCTGGCGCGGAGCCTGCTGCTGCGCGCTGCGCTTGGCTTGGCTTGGCGGCTG	2370
QY	2322	GCTCTGCTGAGGAGCAGGCGGAGGCGCGAGTGGCAACAAGTGGGCAACCGGCGCTACG	2381
Db	2371	GCTCTGCTGAGGAGCAGGCGGAGGCGCGAGTGGCAACAAGTGGGCAACCGGCGCTACG	2430
QY	2382	GCAGAGCTCTTCTGAGCTGCGCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2441
Db	2431	GCAGAGCTCTTCTGAGCTGCGCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2490
QY	2442	GCGCTGCGCACCTGTGATCTGGAAGCTGTGCGCCAGGCGAGTACCGGATCCCTGACCA	2501
Db	2491	GCGCTGCGCACCTGTGATCTGGAAGCTGTGCGCCAGGCGAGTACCGGATCCCTGACCA	2550
QY	2502	TGGAATTTTCTTTTTCGAGGAGCTACTGCTGCGGACCTCGGCGCCCGCCAGAGTCCAGCC	2561
Db	2551	TGGAATTTTCTTTTTCGAGGAGCTACTGCTGCGGACCTCGGCGCCCGCCAGAGTCCAGCC	2610
QY	2562	CCTTGGCCCAACCCCGCTGGACCCAAAGTGTCTGTAGAGAGGSCACCGCCCGGCTGAGT	2621
Db	2611	CCTTGGCCCAACCCCGCTGGACCCAAAGTGTCTGTAGAGAGGSCACCGCCCGGCTGAGT	2670
QY	2622	CCTGGGATATCCGTTTCGACGCTGGAAGAGCGCTTCTGGAAGCGCGCAGCGCCGCTG	2681
Db	2671	CCTGGGATATCCGTTTCGACGCTGGAAGAGCGCTTCTGGAAGCGCGCAGCGCCGCTG	2730
QY	2682	CGGGGCTCAGCTGAGCTTCTACAGGCGCACATCACCGCTTCTGCGGCGCAACCGGG	2741
Db	2731	CGGGGCTCAGCTGAGCTTCTACAGGCGCACATCACCGCTTCTGCGGCGCAACCGGG	2790
QY	2742	GCGGCAAGACCAACCCCTGTCCATCTTGTAGTGGCTCTTCCCAACCCAGTGGTGGCTCT	2801
Db	2791	GCGGCAAGACCAACCCCTGTCCATCTTGTAGTGGCTCTTCCCAACCCAGTGGTGGCTCT	2850
QY	2802	GCTTTCATCTCGGCGCAACGCTGCGCTCCAGATGGCGCGCATCGGCGCCCAACCTGGGC	2861
Db	2851	GCTTTCATCTCGGCGCAACGCTGCGCTCCAGATGGCGCGCATCGGCGCCCAACCTGGGC	2910
QY	2862	GCTGCTCTCAGTCAACAGCTGCTGTTGACATGCTGACCGTGGACGAGCAGCTGCTGTTT	2921

Qy	5082	CTGTATGGCTGGTGCATCACACCGCTCATGTACCCAGCCTCTTCTTCTTCTTCTCCGTGCCC	5141
Db	5131	CTGTATGGCTGGTGCATCACACCGCTCATGTACCCAGCCTCTTCTTCTTCTTCTCCGTGCCC	5190
Qy	5142	AGCACGCCCTATGTGGTCTCACTGTCATAAACCTCTTTATTTGGCATCAATGAAGCATG	5201
Db	5191	AGCACGCCCTATGTGGTCTCACCTGCATAAACCTCTTTATTTGGCATCAATGAAGCATG	5250
Qy	5202	GCCACCTTTGTGCTTGAGCTCTTCTGTATCAGNAGCTGCAGGAGTGAACCGGATCTTG	5261
Db	5251	GCCACCTTTGTGCTTGAGCTCTTCTGTATCAGNAGCTGCAGGAGTGAACCGGATCTTG	5310
Qy	5262	AAACAGGTCTTCCTTATCTTCTCCCCCACTTCTGTCTTGGCCGCGGGGCTCATTTGACATGGTG	5321
Db	5311	AAACAGGTCTTCCTTATCTTCTCCCCCACTTCTGTCTTGGCCGCGGGGCTCATTTGACATGGTG	5370
Qy	5322	CGGAACAGGCGCATGGCTGATGCGCTTTGAGCGCTTGGGAGACAGGCAAGTTCCAGTCAACC	5381
Db	5371	CGGAACAGGCGCATGGCTGATGCGCTTTGAGCGCTTGGGAGACAGGCAAGTTCCAGTCAACC	5430
Qy	5382	CTGCGCTGGAGGTGGTTCGGCAAGAACCTCTTGGCCATGCTGATACAGGGGCCCCCTTTC	5441
Db	5431	CTGCGCTGGAGGTGGTTCGGCAAGAACCTCTTGGCCATGCTGATACAGGGGCCCCCTTTC	5490
Qy	5442	CTTCTCTTTACACTACTGCTGCAGCACCGCAAGCCAACTCTCTGCCACAGCCACGGGTGAGG	5501
Db	5491	CTTCTCTTTACACTACTGCTGCAGCACCGCAAGCCAACTCTCTGCCACAGCCACGGGTGAGG	5550
Qy	5502	TCTCTGCCACTCTCTGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGCTGGTC	5561
Db	5551	TCTCTGCCACTCTCTGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGCTGGTC	5610
Qy	5562	CAAGGAGCACCCAGGGGGATGTGTTGGTCTGAGGAACTTTGACCAAGGTATACCGTGGG	5621
Db	5611	CAAGGAGCACCCAGGGGGATGTGTTGGTCTGAGGAACTTTGACCAAGGTATACCGTGGG	5670
Qy	5622	CAGAGGATGCCAGCTGTTTGACCGCTGTGCTCGGGGATTCCTCCCTGCTGAGTGTGTTTGGG	5681
Db	5671	CAGAGGATGCCAGCTGTTTGACCGCTGTGCTCGGGGATTCCTCCCTGCTGAGTGTGTTTGGG	5730
Qy	5682	CTGCTGGGTGTGAATGGAGCAGGGAAGACGTCACGTTTCGCATGTGTGACGGGGGACACA	5741
Db	5731	CTGCTGGGTGTGAATGGAGCAGGGAAGACGTCACGTTTCGCATGTGTGACGGGGGACACA	5790
Qy	5742	TTGCCACAGCGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAAACCCAGTGCT	5801
Db	5791	TTGCCACAGCGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAAACCCAGTGCT	5850
Qy	5802	GCGCACCTCAGCATGGGATATCGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5861
Db	5851	GCGCACCTCAGCATGGGATATCGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5910
Qy	5862	GCGCAGCACCTGGAGCTGCTTGGCGGCTGCGCGGTGTCCCGAGGCCACAGGTTGCCAG	5921
Db	5911	GCGCAGCACCTGGAGCTGCTTGGCGGCTGCGCGGTGTCCCGAGGCCACAGGTTGCCAG	5970
Qy	5922	ACCGCTGGCTCAGCGCTGGCGGCTCTGGGACTCTCATGTGTACGCAGACCGGCGCTGAGGC	5981
Db	5971	ACCGCTGGCTGGCGGCTGGCGGCTCTGGGACTCTCATGTGTACGCAGACCGGCGCTGAGGC	6030
Qy	5982	ACCTACAGCGGAGGGAACAAACGCAAGCTGGCAGCGCCCTGGCGCTGTTGGGACCCA	6041
Db	6031	ACCTACAGCGGAGGGAACAAACGCAAGCTGGCAGCGCCCTGGCGCTGTTGGGACCCA	6090
Qy	6042	GCCGTGGTGTCTTGGACGAGCCACCAAGGACATGGACCCCAACGCGCGCGGCTTCTT	6101
Db	6091	GCCGTGGTGTCTTGGACGAGCCACCAAGGACATGGACCCCAACGCGCGCGGCTTCTT	6150
Qy	6102	TGGAAACAGCCCTTTTGGCGGTGGTGGGAGGGCGCTTCACTGATGCTCACTTCCCATAGC	6161
Db	6151	TGGAAACAGCCCTTTTGGCGGTGGTGGGAGGGCGCTTCACTGATGCTCACTTCCCATAGC	6210

Qy	6162	ATGAGAGAGTGTGAAGCGCTCTCTCTCGCGCTAGCATTATGTGTGAATGGCGGTTCCGC	6222
Db	6211	ATGAGAGAGTGTGAAGCGCTCTCTCTCGCGCTAGCATTATGTGTGAATGGCGGTTCCGC	6270
Qy	6222	TGCCTGGCAGCCGCAACATCTCAAGGCAGATTTCGGCGGGTCAACACTGACCTGACCCCTG	6281
Db	6271	TGCCTGGCAGCCGCAACATCTCAAGGGC-----	6300
Qy	6282	CGGTTGCCCGCCGCAAGAGTCCCAAGCGGAGCGGCTTCTGTGGCGGCGAGTTCCCTTGGG	6341
Db	6301	-----AGGTCCAGCCGCGAGCGGCTTCTGTGGCGGCGAGTTCCCTTGGG	6345
Qy	6342	TCGAGAGTGCAGGAGCACAATGAGGCGCGCTTCGCTTCCAGTGCCTGCGCGGAGGCGC	6401
Db	6346	TCGAGAGTGCAGGAGCACAATGAGGCGCGCTTCGCTTCCAGTGCCTGCGCGGAGGCGC	6405
Qy	6402	TGGCCCTTGGCGCGGCTTCTTGGAGAGCTGGCGGTCGACGGCCGAGAGCACGGCGTGGAG	6461
Db	6406	TGGCCCTTGGCGCGGCTTCTTGGAGAGCTGGCGGTCGACGGCCGAGAGCACGGCGTGGAG	6465
Qy	6462	GACTTTTCCGTGAGCCAGACGATGCTGGAGAGGTATTTCTGTACTTCTCCAAAGGACCAG	6521
Db	6466	GACTTTTCCGTGAGCCAGACGATGCTGGAGAGGTATTTCTGTACTTCTCCAAAGGACCAG	6525
Qy	6522	GGGAAGGACGAGGACACCCGAAGAGCAGAAGGAGGAGTGGAGTGGACCCCGCGCCA	6581
Db	6526	GGGAAGGACGAGGACACCCGAAGAGCAGAAGGAGGAGTGGAGTGGACCCCGCGCCA	6585
Qy	6582	GGCCTGAGCACCACCCAAACCGCTCAGCCAGTTCTCTGATGACCTAGCCTGCGAGACT	6641
Db	6586	GGCCTGAGCACCACCCAAACCGCTCAGCCAGTTCTCTGATGACCTAGCCTGCGAGACT	6645
Qy	6642	GTGCTC 6647	
Db	6646	GTGCTC 6651	
RESULT 11			
AAD05626/c			
ID	AAD05626 standard; cDNA; 5811 BP.		
XX	AAD05626;		
XX	17-JUL-2001 (first entry)		
DT	Human secreted protein-encoding gene 11 cDNA clone HWBF57, SEQ ID NO:58.		
XX	Human; secreted protein; proliferative disorder; cancer; tumour;		
KW	foetal abnormality; developmental abnormality; haematopoietic disorder;		
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;		
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;		
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;		
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;		
KW	cardiovascular disorder; angiogenic disorder; kidney disorder;		
KW	gastrointestinal disorder; pregnancy-related disorder;		
KW	endocrine disorder; infection; wound healing; vulvuary; cell culture;		
KW	chemotaxis; food additive; gene therapy; binding partner identification;		
ss.			
XX	Homo sapiens.		
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	3300..3413	
FT		/*tag= a	
FT		/product= "Human secreted protein"	
FT		/note= "CDS does not include start codon"	
FT		/partial	
FT	sig_peptide	3300..3347	
FT		/*tag= b	
FT	mat_peptide	3348..3410	
FT		/*tag= c	
FT		/product= "Human mature secreted protein"	
XX			

PN WO200134627-A1.
XX 17-MAY-2001.
XX 08-NOV-2000; 2000WO-US030628.
XX 12-NOV-1999; 99US-0164744P.
PR 30-JUN-2000; 2000US-0215140P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatoulis GA, Baker KP, Young PE;
XX WPI; 2001-316491/33.
DR P-PSDB; AAE01817.
XX New nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX Claim 1; Page 457-458; 567pp; English.
XX AAD05579-AAD05658 represent cDNAs corresponding to 28 human secreted
CC protein genes and AAE01770-AAE01849 represent the proteins they encode.
CC AAE01850-AAE01860 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 28 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angioneurotic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein-encoding cDNA of the
XX invention
XX Sequence 5811 BP; 1068 A; 1890 C; 1931 G; 1006 T; 0 U; 16 Other;
Query Match 81.3%; Score 5529; DB 4; Length 5811;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 5709; Conservative 12; Mismatches 40; Indels 125; Gaps 4;
QY 921 CTCACCTGTTACGAGCTAGTACCTGATGAGCTGTGGGGCAGAGCCAGAAATCGCC 980
DB 5801 CTTCACTGGTACGANGCTAGTACCTGATGAGCTGTGGGGCAGAGCCAGAAATCGCC 5742
QY 981 CTGCAGACAGCAGCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 1040
DB 5741 CTGCAGACAGCAGCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 5682
QY 1041 CCGCTGTCCCGCTGCTCTGGAGACGCTGTAAGCCTCTGTATCTCTCGGAACTACTCTTT 1100
DB 5681 CCGCTGTCCCGCTGCTCTGGAGACGCTGTAAGCCTCTGTATCTCTCGGAACTACTCTTT 5622
QY 1101 GCACAGATACACCTTTTACCGGAAGCTCATGGCCCGAGTGAACCGGACCTTCGAGGAG 1160
DB 5621 GCACAGATACACCTTTTACCGGAAGCTCATGGCCCGAGTGAACCGGACCTTCGAGGAG 5562

QY 1161 CTCACCTGCTGAGGAGTGTCCGGAGAGTGTGGAGATGCTGGGACCCCGGATCTTTCACC 1220
DB 5561 CTCACCTGCTGAGGAGTGTCCGGAGAGTGTGGAGATGCTGGGACCCCGGATCTTTCACC 5502
QY 1221 TTCAATGAACGACAGTTCATGTCGTCATGTCGAGCGGCTCTGTCAGATGACAGATGAA 1280
DB 5501 TTYATGAACGACAGTTCATGTCGTCATGTCGAGCGGCTCTGTCAGATGACAGATGAA 5442
QY 1281 GGAAGAGGAGCCAGCCAGACCTGGAGGCGGAGACACATGGAGGCGCTCCGATCTCTTCG 1340
DB 5441 GGAAGAGGAGCCAGACCTGGAGGCGGAGACACATGGAGGCGCTCCGATCTCTTCG 5382
QY 1341 GACCTCTGGAGCGGTGTACAGCTGGCAGACGACACACGCTGATGTGGGACCCCTGGTG 1400
DB 5381 GACCTCTGGAGCGGTGTACAGCTGGCAGACGACACACGCTGATGTGGGACCCCTGGTG 5322
QY 1401 GGCAAGCTGGGCGGAGTGAACGAGTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1460
DB 5321 GGCAAGCTGGGCGGAGTGAACGAGTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5262
QY 1461 GAGCAGCCCTGTGTGTCGGGCGCTGCAACTGCTCGCGGAACATCGATTCTGGGCGGCG 1520
DB 5261 GAGCAGCCCTGTGTGTCGGGCGCTGCAACTGCTCGCGGAACATCGATTCTGGGCGGCG 5202
QY 1521 GTCGCTCTTCTTGGGACCTGAGGACTCTTTCAGACCCCGACAGACACCCAGACCTG 1580
DB 5201 GTCGCTCTTCTTGGGACCTGAGGACTCTTTCAGACCCCGACAGACACCCAGACCTG 5142
QY 1581 GGCCCGGCGCACGTCGGGATCAAAATCGCATGATGACATGATGATGATGATGATGATGAT 1640
DB 5141 GGCCCGGCG-CACGTGCGCATCAAAATCGCATGATGATGATGATGATGATGATGATGAT 5083
QY 1641 AAGATAGGAGACAGTGTGGGACCTTGGGACCTTGGGACCTTGGGACCTTGGGACCTG 1700
DB 5082 AAGATAGGAGACAGTGTGGGACCTTGGGACCTTGGGACCTTGGGACCTTGGGACCTG 5023
QY 1701 TACGTGTGGGCGGCTTGTGTACTCTGCAAGACCTGTGTGGAGCGTGTGACGCGTGTG 1760
DB 5022 TACGTGTGGGCGGCTTGTGTACTCTGCAAGACCTGTGTGGAGCGTGTGACGCGTGTG 4963
QY 1761 CTACAGCGGCGCAACCCCGGCGGCTTCTTACCTGACAGATGATGATGATGATGATGATGAT 1820
DB 4962 CTACAGCGGCGCAACCCCGGCGGCTTCTTACCTGACAGATGATGATGATGATGATGATGAT 4903
QY 1821 GTGACAGCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1880
DB 4902 GTGACAGCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4843
QY 1881 TGGATCTACTCCGTGACACCTGACAGTGAAGCCCTGTGTGGGAGAGAGAGAGAGAGAGAG 1940
DB 4842 TGGATCTACTCCGTGACACCTGACAGTGAAGCCCTGTGTGGGAGAGAGAGAGAGAGAGAG 4783
QY 1941 CGGAGACACATGTCGCGCATGGGCTCAGCGCGCGGTGTCTGTGGTAGGTGTGTGTCTCTC 2000
DB 4782 CGGAGACACATGTCGCGCATGGGCTCAGCGCGCGGTGTCTGTGGTAGGTGTGTGTCTCTC 4723
QY 2001 AGCTGCTCGGCGCTTCTGCTCAGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2060
DB 4722 AGCTGCTCGGCGCTTCTGCTCAGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4663
QY 2061 GACATCTCTCCCTACAGCACCCCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2120
DB 4662 GACATCTCTCCCTACAGCACCCCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4603
QY 2121 GCCACGCTGACCCAGAGCTTCTGCTCAGCGCGCTTCTTCTCCCGCGGCAACCTGGCTGCG 2180
DB 4602 GCCACGCTGACCCAGAGCTTCTGCTCAGCGCGCTTCTTCTCCCGCGGCAACCTGGCTGCG 4543
QY 2181 GCCTGCGGCGGCTGCGGCTTCTTCT 2240
DB 4542 GCCTGCGGCGGCTGCGGCTTCTTCT 4483

QY 6619 ATGACCTTAGCTGCGGAGCTGTGCTCTGAGCTCCCTCCCTGCGGGCGCGGGA 6678
DB 224 ATGACCTTAGCTGCGGAGCTGTGCTCTGAGCTCCCTCCCTGCGGGCGCGGGA 165
QY 6679 GGCCTTGGGAATGCAAGCAAGGTAGCTGTGCTAGGAGCCCTGGAAGCTGCGCAG 6738
DB 164 GGCCTTGGGAATGCAAGCAAGGTAGCTGTGCTAGGAGCCCTGGAAGCTGCGCAG 105
QY 6739 AGGGCTGTGCTGCGGAGCAATTAAGAGAGCTGGAGAGCGCTGTGTAATA 6798
DB 104 AGGGCTGTGCTGCGGAGCAATTAAGAGAGCTGGAGAGCGCTGTGTAATA 45
QY 6799 AAAAAA 6804
DB 44 AAAAAA 39

RESULT 12
ADA40555/c
ID ADA40555 standard; cDNA; 5811 BP.
XX
AC ADA40555;
XX
DT 20-NOV-2003 (first entry)
XX
XX Human secreted protein encoding cDNA.
XX
XX Human; secreted protein; cancer; hyperproliferative disorder;
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
KW anaemia; allergic reaction; asthma; cardiovascular disorder;
KW wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;
KW antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory;
KW vulnerable; cardiac; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO2002102993-A2.
XX
XX 27-DEC-2002.
XX
XX 19-MAR-2002; 2002WO-US008123.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX
XX 19-JUL-2001; 2001US-0306171P.
XX
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-175238/17.
XX
XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.
XX
XX Claim 9; SEQ ID NO 937; 3205pp; English.
XX
XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,

CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5811 BP; 1068 A; 1890 C; 1831 G; 1006 T; 0 U; 16 Other;
Query Match 81.3%; Score 5529; DB 8; Length 5811;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 5709; Conservative 12; Mismatches 40; Indels 125; Gaps 4;
QY 921 CTCAACTGTGTAGAGGCTAGTGACCTGATGAGCTGGTGGGGAGGAGCCAGATCCGCC 980
DB 5801 CTTCACTGTGTAGGANGCTAGTGACCTGATGAGCTGGTGGGGAGGAGCCAGATCCGCC 5742
QY 981 CTGCCAGACAGAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 1040
DB 5741 CTGCCAGACAGAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 5682
QY 1041 CCGCTGTCCCGCTGCTCTGGAGAGCCCTGAAAGCTCTGATCTCTGGGAGAGTACTCTTT 1100
DB 5681 CCGCTGTCCCGCTGCTCTGGAGAGCCCTGAAAGCTCTGATCTCTGGGAGAGTACTCTTT 5622
QY 1101 GCACAGATACACCTTTTACC CGGAAGCTCATG GCGCCAGGTGAACCGGACCTTCGAGGAG 1160
DB 5621 GCACAGATACACCTTTTACC CGGAAGCTCATG GCGCCAGGTGAACCGGACCTTCGAGGAG 5562
QY 1161 CTCACCTCTGAGGGATGTCGGGAGGTGCGGAGATGCTGGGACCCCGGATCTTCACC 1220
DB 5561 CTCACCTCTGCGGGATGTCGGGAGGTGCGGAGATGCTGGGACCCCGGATCTTCACC 5502
QY 1221 TTCTGAACGACAGATTTCCAAATGTGGCCATGCTGACGCGCTCTCTGCAGATGACGAGTGA 1280
DB 5501 TTCTGAACGACAGATTTCCAAATGTGGCCATGCTGACGCGCTCTCTGCAGATGACGAGTGA 5442
QY 1281 GGAAGAGGCGAGCCAGACCTGAGAGCCCGGACCAATGAGAGCCCTGCGATCCTTTCTG 1340
DB 5441 GGAAGAGGCGAGCCAGACCTGAGAGCCCGGACCAATGAGAGCCCTGCGATCCTTTCTG 5382
QY 1341 GACCTGGGAGCGGTGGCTACAGCTGGCAGAGCGCACACGCTGATGTGGGACCTGGTG 1400
DB 5381 GACCTGGGAGCGGTGGCTACAGCTGGCAGAGCGCACACGCTGATGTGGGACCTGGTG 5322
QY 1401 GGCACGCTGGGCGCGAGTGACGAGTGCCTGTCTCTTGGCAAGCTGGAGGCGGACCCCTCA 1460
DB 5321 GGCACGCTGGGCGCGAGTGACGAGTGCCTGTCTCTTGGCAAGCTGGAGGCGGACCCCTCA 5262
QY 1461 GAGGAGCCCTGTGTGTCGGGCGCTTGAACTGTCTCGGGAACATCGATTTCTGGGCGGC 1520
DB 5261 GAGGAGCCCTGTGTGTCGGGCGCTTGAACTGTCTCGGGAACATCGATTTCTGGGCGGC 5202
QY 1521 GTCTGCTTTCTTGGGACCTGAGGACTTTCAGAGCCCGCCACAGAGCCACCCAGACCTG 1580
DB 5201 GTCTGCTTTCTTGGGACCTGAGGACTTTCAGAGCCCGCCACAGAGCCACCCAGACCTG 5142
QY 1581 GGGCCCGGCGAGCTGCGCATCAAAATCCGCATGGAATTCGATGTGTGTCAGAGGACCAAT 1640
DB 5141 GGGCCCGG-CACGTGGCATCAAAATCCGCATGGAATTCGATGTGTGTCAGAGGACCAAT 5083
QY 1641 AAGATCAGGAGCAGGTTTGGGACCTTGCCCGAGCCCGGAGCCCTCTGACCGACCTGGC 1700
DB 5082 AAGATCAGGAGCAGGTTTGGGACCTTGCCCGAGCCCGGAGCCCTCTGACCGACCTGGC 5023

QY 1701 TACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCGTCCGCGTG 1760
DB 5022 TACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCGTCCGCGTG 4963
QY 1761 CTCAGCGCGCCAAACCCCGGGCGGCTCTACCTGCAGACAGATGCCCTATCCGTGCTAT 1820
DB 4962 CTCAGCGCGCCAAACCCCGGGCGGCTCTACCTGCAGACAGATGCCCTATCCGTGCTAT 4903
QY 1821 GTGAGACAGCTGTTCCTCGGTGTGTAGCCGGTCCGTCCGCTCTTCTGAACGTGGCC 1880
DB 4902 GTGAGACAGCTGTTCCTCGGTGTGTAGCCGGTCCGTCCGCTCTTCTGAACGTGGCC 4843
QY 1881 TGGATCTACTCCGTGACACTGACAGTGAAGCCGCTGGTCCGGGAGAGACCGGCTG 1940
DB 4842 TGGATCTACTCCGTGACACTGACAGTGAAGCCGCTGGTCCGGGAGAGACCGGCTG 4783
QY 1941 CGGACACATGCGCGCCATGGGGCTCAGCCGCGCGTCTGTGGCTAGGCTGGTTCCTC 2000
DB 4782 CGGACACATGCGCGCCATGGGGCTCAGCCGCGCGTCTGTGGCTAGGCTGGTTCCTC 4723
QY 2001 AGCTGCTCGGGCCCTTCCTGCTCAGCCCGCGCTGCTGTGGTCTCAAGCTGGGG 2060
DB 4722 AGCTGCTCGGGCCCTTCCTGCTCAGCCCGCGCTGCTGTGGTCTCAAGCTGGGG 4663
QY 2061 GACATCTCTCCCTACAGCCACCGGGCGTGTCTCTCTGTTCTTGGCAGCCTTCCGCGTG 2120
DB 4662 GACATCTCTCCCTACAGCCACCGGGCGTGTCTCTCTGTTCTTGGCAGCCTTCCGCGTG 4603
QY 2121 GCCAGGTGACCCAGAGCTTCCTGCTCAGCGCTTCCTTCGCCGCGCAACCTGGCTGCG 2180
DB 4602 GCCAGGTGACCCAGAGCTTCCTGCTCAGCGCTTCCTTCGCCGCGCAACCTGGCTGCG 4543
QY 2181 GCCTCGCGCGCTGCGCTACTTCTCCCTCTACCTGCGCTACGTGCTGTGTGGCTGG 2240
DB 4542 GCCTCGCGCGCTGCGCTACTTCTCCCTCTACCTGCGCTACGTGCTGTGTGGCTGG 4483
QY 2241 CGGACACCGGCTGCGCGGGTGGCGCGTGGCGCGAGCTGTGTGCCCGTGGCC-TT 2299
DB 4482 CGGACACCGGCTGCGCGGGTGGCGCGTGGCGCGAGCTGTGTGCCCGTGGCC-TT 4423
QY 2300 CGGCTTCGGCTGCG-AGAGCTTGGCTCTGTGGAGGACAGGGCGGCGCAGTGGC 2358
DB 4422 CGGCTTCGGCTGCGGAAAGCTGCGCTCTGTGGAGGACAGGGCGGCGCAGTGGC 4363
QY 2359 ACAAGCTGGGCAACCGGCTACCGGACAGCTCTTCAGCGCTGCGCCAGAGTCTGTGGCTTC 2418
DB 4362 ACAAGCTGGGCAACCGGCTACCGGACAGCTCTTCAGCGCTGCGCCAGAGTCTGTGGCTTC 4303
QY 2419 TGCTGTGGACCGCGCGCTCTACCGCTTCGCCACCTGGTACTGGAAAGCTGTGCCAG 2478
DB 4302 TGCTGTGGACCGCGCGCTCTACCGCTTCGCCACCTGGTACTGGAAAGCTGTGCCAG 4243
QY 2479 GCCAGTACGGATCCCTGAAACATGGAAATTTCTTTTCGAGAGCTACTGTGGCGGAC 2538
DB 4242 GCCAGTACGGATCCCTGAAACATGGAAATTTCTTTTCGAGAGCTACTGTGGCGGAC 4183
QY 2539 CTCGCGCCCCCAAGAGTCCAGCCCTTCCGCCACCCCGCTGACCCAAAGGTGTGGTAG 2598
DB 4182 CTCGCGCCCCCAAGAGTCCAGCCCTTTCGCCACCCCGCTGACCCAAAGGTGTGGTAG 4123
QY 2599 AAGAGGACCCCGCGCTGAGTCTGGCGCTATCCGTTTCGAGCGCTGGAGAGCGCTTTC 2658
DB 4122 AAGAGGACCCCGCGCTGAGTCTGGCGCTCTCGTTTCGAGCGCTGGAGAGCGCTTTC 4063
QY 2659 CTGGAAGCCCGAGCAGCCCTTGGGGGGCTGAGCTTGAATTCTTACCGGGCCACATCA 2718
DB 4062 CTGGAAGCCCGAGCAGCCCTTGGGGGGCTGAGCTTGAATTCTTACCGGGCCACATCA 4003
QY 2719 CCGCTTCTGGGCCACAAAGGGCGCGAGACCAACCCCTGCTCCATCTTGTAGTGGCC 2778
DB 4002 CCGCTTCTGGGCCACAAAGGGCGCGAGACCAACCCCTGCTCCATCTTGTAGTGGCC 3943
QY 2779 TCTTCCCAACCGAGTGGTGGCTCTGCTTCTATCTCTGGGCGACGACGTCCGCTCCAGCATGG 2838

DB 3942 TCTTCCCAACCGAGTGGTGGCTCTGCTTCTATCTGGGCGACAGCTCGCTCCAGCATGG 3883
QY 2839 CCGCATCCGGCCCGCCACCTGGCGCTCTGCTCTCAGTACAAAGTGCTGTGTTGACATGCTGA 2898
DB 3882 CCGCATCCGGCCCGCCACCTGGCGCTCTGCTCTCAGTACAAAGTGCTGTGTTGACATGCTGA 3823
QY 2899 CCGTGGACGACACGTCTGTGTTCTATGGCGGCTGAAGGGTCTGAGTCCCGCTGTAGTGG 2958
DB 3822 CCGTGGACGACACGTCTGTGTTCTATGGCGGCTGAAGGGTCTGAGTCCCGCTGTAGTGG 3763
QY 2959 GCCCGAGACGACCGCTGTGCTGCAAGTGTGGGGCTGGTCTCCAGCAGAGTGTGCAGA 3018
DB 3762 GCCCGAGACGACCGCTGTGCTGAGGATGTGGGGCTGGTCTCCAGCAGAGTGTGCAGA 3703
QY 3019 CTGCGCACCTCTCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTGTGGGCG 3078
DB 3702 CTGCGCACCTCTCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTGTGGGCG 3643
QY 3079 GCTCCCAAGTGTATCTCTGACGAGCTACCGGTGGCGTGGATCTCTGCTTCCCGCCGCG 3138
DB 3642 KCKCCCAAGTGTATCTCTGACGAGCTACCGGTGGCGTGGATCTCTGCTTCCCGCCGCG 3583
QY 3139 GTATTTGGGAGCTGCTGCTCAATACCGAGAGGTTCGACGCTGATCTCTCCACCCACC 3198
DB 3582 GTATTTGGGAGCTGCTGCTCAATACCGAGAGGTTCGACGCTGATCTCTCCACCCACC 3523
QY 3199 ACCTGGATGAGGACAGAGCTGCTGGAGACCTGTGGCTGTGGTGGAGTGGCGCTTGT 3258
DB 3522 ACCTGGATGAGGACAGAGCTGCTGGAGACCTGTGGCGCTGTGGCAGTGGCGCTTGT 3463
QY 3259 GCTGCTGTGGCTCCCACTCTTCTGCGCCCTCACCTGGGCTCCGGCTACTACTGAGCGC 3318
DB 3462 GCTGCTGTGGCTCCCACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACTACTGAGCGC 3403
QY 3319 TGGTGAAGGCCCGCTGCTGCTGACACCAATGAGAGGCTGACACTGACATGAGAGGCA 3378
DB 3402 TGGTGAAGGCCCGCTGCTGCTGACACCAATGAGAGGCTGACACTGACATGAGAGGCA 3343
QY 3379 GTGTGGACACCGAGCAGGAAAGAAATGACGCCAGGCGAGAGTCCGCACTCTC 3438
DB 3342 GTGTGGACACCGAGCAGGAAAGAAATGACGCCAGGCGAGAGTCCGCACTCTC 3283
QY 3439 AGCTGCTGGCCCTGGTACAGCACTGGGTGCGCGGCGCACGCTGTGGTGGAGAGTCCGAC 3498
DB 3282 AGCTGCTGGCCCTGGTACAGCACTGGGTGCGCGGCGCACGCTGGTGGAGAGTCCGAC 3223
QY 3499 ACAGCTGTGTGCTGCTGCTTACAGCGGTGCCATGACGCGAGCTTCCGCACTCT 3558
DB 3222 ACAGCTGTGTGCTGCTGCTTACAGCGGTGCCATGACGCGAGCTTCCGCACTCT 3163
QY 3559 TCCGAGAGCTAGACACCGGCTGGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3618
DB 3162 TCCGAGAGCTAGACACCGGCTGGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3103
QY 3619 CCAGCTCGAGGAGATCTTCTGAAAGTGTGGAGGAGTGTGTCGCGACACAGATATGG 3678
DB 3102 CCAGCTCGAGGAGATCTTCTGAAAGTGTGGAGGAGTGTGTCGCGACACAGATATGG 3043
QY 3679 AGGATGAGCTGCGGCGAGCCTATGCAAGGCAATTTGTGGCTTGAACCTTACCCCTGC 3738
DB 3042 AGGATGAGCTGCGGCGAGCCTATGCAAGGCAATTTGTGGCTTGAACCTTACCCCTGC 2983
QY 3739 GGCTCAAGATGCCGCCACAGAGACAGGCTGGAGAACGGGGAACAGCTGGGTGAGCCC 3798
DB 2982 GGCTCAAGATGCCGCCACAGAGACAGGCTGGAGAACGGGGAACAGCTGGGTGAGCCC 2923
QY 3799 CAGAGATGACAGGGCTCTGGGCGCAGACCGCTGGGCGGGGTACAGGGCTGGGCACTGA 3858
DB 2922 CAGAGATGACAGGGCTCTGGGCGCAGACCGCTGGGCGGGGTACAGGGCTGGGCACTGA 2863
QY 3859 CCGGCCAGAGCTCCAGGCGCTTCTCAAGCGCTTCTGCTTGGCGCGCGAGCGGCC 3918

QY 6079 ACCCAGCGCGCGCGCTTCTTTGGAACAGCCTTTTGGCCGTGTGTGGGAGGCGCGTT 6138
DB 764 ACCCAGCGCGCGCGCTTCTTTGGAACAGCCTTTTGGCCGTGTGTGGGAGGCGCGTT 705
QY 6139 CAGTGATGCTACCTCCCATAGCATGAGAGGTGTGAAGCGCTCTGCTCGCGCCTAGCCA 6198
DB 704 CAGTGATGCTACCTCCCATAGCATGAGAGGTGTGAAGCGCTCTGCTCGCGCCTAGCCA 645
QY 6199 TCATGGTGAATGGCGGTTCCGCTGCTGGGAGCGCCGCAACATCTCAAGGCGAGATTG 6258
DB 644 TCATGGTGAATGGCGGTTCCGCTGCTGGGAGCGCCGCAACATCTCAAGGCGAGATTG 585
QY 6259 CGCGGGGTACACATGACCTGCGGGTGCCTGCGGAGGTCCTGAGGCGCGGCGCT 6318
DB 584 CGCGGGGTACACATGACCTGCGGGTGCCTGCGGAGGTCCTGAGGCGCGGCGCT 525
QY 6319 TCGTGGCGCGGAGTTCCTGCGGAGTGCAGGAGGACATGAGGCGCGCTGCGCT 6378
DB 524 TCGTGGCGCGGAGTTCCTGCGGAGTGCAGGAGGACATGAGGCGCGCTGCGCT 465
QY 6379 TCCAGCTGCCGCGGAGGCGCTGCGCCCTGCGCGGCTCTTTTGGAGAGCTGGCGGTGC 6438
DB 464 TCCAGCTGCCGCGGAGGCGCTGCGCCCTGCGCGGCTCTTTTGGAGAGCTGGCGGTGC 405
QY 6439 ACGGCGAGAGCAGCGCGTGGAGGACTTTTCGTGAGCCAGACGATGCTGAGGAGGTAT 6498
DB 404 ACGGCGAGAGCAGCGCGTGGAGGACTTTTCGTGAGCCAGACGATGCTGAGGAGGTAT 345
QY 6499 TCTGTACTTCTCCAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6558
DB 344 TCTGTACTTCTCCAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 285
QY 6559 GAGTGGAGTGGACCCCGCGGAGGCTGCGAGCACCACCAAGCGCTCAGCGATTCTCTCG 6618
DB 284 GAGTGGAGTGGACCCCGCGGAGGCTGCGAGCACCACCAAGCGCTCAGCGATTCTCTCG 225
QY 6619 ATGACCTTAGCAGTGCAGAGCTGTGCTGAGGCTTCCCTCCCTGCGGGGCGCGGGGA 6678
DB 224 ATGACCTTAGCAGTGCAGAGCTGTGCTGAGGCTTCCCTCCCTGCGGGGCGCGGGGA 165
QY 6679 GGCCCTGGGATGCAAGGCAAGGTAGAGTGCCTTAGGAGCCCTGAGCTCAGGCTGGCAG 6738
DB 164 GGCCCTGGGATGCAAGGCAAGGTAGAGTGCCTTAGGAGCCCTGAGCTCAGGCTGGCAG 105
QY 6739 AGGGGCTGGTCCCTGGAGAAATAAAGAGAGGCTGAGAGAGCCCTGCTGGTGAATA 6798
DB 104 AGGGGCTGGTCCCTGGAGAAATAAAGAGAGGCTGAGAGAGCCCTGCTGGTGAATA 45
QY 6799 AAAAAA 6804
DB 44 AAAAAA 39

RESULT 13

ID ADA56706/c
AC ADA56706 standard; DNA; 5811 BP.
XX ADA56706;
XX

DT 20-NOV-2003 (first entry)

DE Gene encoding human secreted protein #578.

XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
KW cytosolic; cerebroprotective; neuroprotective; nootropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology; ds;
KW gene.

XX Homo sapiens.
OS WO2002102994-A2.
PN 27-DEC-2002.
XX

PD 19-MAR-2002; 2002WO-US008278.
XX

PF 21-MAR-2001; 2001US-0277340P.
XX

PR 19-JUL-2001; 2001US-0306171P.
XX

PP 13-NOV-2001; 2001US-0331287P.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Ruben SM;
XX

XX WPI; 2003-167512/16.
XX

DR P-PSDB; ADA57599.
XX

XX New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX

PS Claim 21; SEQ ID NO 895; 1754pp; English.
XX

XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that binds to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to a gene encoding one of the polypeptide of the invention. Note: The
CC sequence data for this patent did form part of the printed specification,
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 5811 BP; 1068 A; 1890 C; 1831 G; 1006 T; 0 U; 16 Other;

XX Query Match 81.3%; Score 5529; DB 10; Length 5811;

XX Best Local Similarity 97.0%; Pred. No. 0;

XX Matches 5709; Conservative 12; Mismatches 40; Indels 125; Gaps 4;

QY 921 CTCACCTGGTACGAGGCTAGTACCTGATGAGCTGTGGGCGAGGCCAGATCCGCC 980

DB 5801 CTTTCACCTGGTACGAGGCTAGTACCTGATGAGCTGTGGGCGAGGCCAGATCCGCC 5742

QY 981 CTGCAGACAGCAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCTCGACAGCCAC 1040

DB 5741 CTGCAGACAGCAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCTCGACAGCCAC 5682

QY 1041 CCGCTGTCCCGCTGTCTCTGGAGAGCGCTGAAGCCTCTGATCTCTGGGAAGCTACTCTTT 1100

Db 5681 CGCTGTCCCGCTGCTCGAGAGCGCTGAAGCCTCTGATCCTCGGAAGCTACTCTTT 5622
QY 1101 GCACCAAGATACACTTTTACCCGGAAGCTCATGGCCAGAGTGAAACGGACCTTCGAGGAG 1160
Db 5621 GCACCAAGATACACTTTTACCCGGAAGCTCATGGCCAGAGTGAAACGGACCTTCGAGGAG 5562
QY 1161 CTCACCTCTGTGAGGAGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACC 1220
Db 5561 CTCACCTCTGTGAGGAGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACC 5502
QY 1221 TTCAATGAACGACATGTTCCAAATGTGGCCATGCTCAGCGGCTCCTCGAGATGCAAGATGAA 1280
Db 5501 TTVAATGAACGACATGTTCCAAATGTGGCCATGCTCAGCGGCTCCTCGAGATGCAAGATGAA 5442
QY 1281 GGAAGAGGAGCGCCAGACTGAGAGCGGGGACCAATGAGAGCCCTGCGATCCTTTCTG 1340
Db 5441 GGAAGAGGAGCGCCAGACTGAGAGCGGGGACCAATGAGAGCCCTGCGATCCTTTCTG 5382
QY 1341 GACCTGGAGCGGTGGCTACAGCTGGCAGGAGCGACACGCTGATGTGGGACCTGGTG 1400
Db 5381 GACCTGGAGCGGTGGCTACAGCTGGCAGGAGCGACACGCTGATGTGGGACCTGGTG 5322
QY 1401 GGCACTGTGGCCGAGTGACGAGTGCTCTCTTGGACAAGCTGGAGCGGACCTCTCA 1460
Db 5321 GGCACTGTGGCCGAGTGACGAGTGCTCTCTTGGACAAGCTGGAGCGGACCTCTCA 5262
QY 1461 GAGGAGCCCTGTGTGTCGGGGCCCTGCAACTGCTGTGCGGAAATGATTCGGGGCGGC 1520
Db 5261 GAGGAGCCCTGTGTGTCGGGGCCCTGCAACTGCTGTGCGGAAATGATTCGGGGCGGC 5202
QY 1521 GTCTGTCTTGTGGACCTGAGGACTCTTACAGACCCACAGAGCACCAACCCAGACCTG 1580
Db 5201 GTCTGTCTTGTGGACCTGAGGACTCTTACAGACCCACAGAGCACCAACCCAGACCTG 5142
QY 1581 GGGCCGGCCACGTGGGCATCAAAATCCGCATGGACATTTGACGTGTGTACAGAGCAAT 1640
Db 5141 GGGCCCGG - CACGTGGCATCAAAATCCGCATGGACATTTGACGTGTGTACAGAGCAAT 5083
QY 1641 AAGATCAGGACAGGTTTGGGACCTTGGCCAGCCAGCCGCGGACCCCTGACGACCTGGCC 1700
Db 5082 AAGATCAGGACAGGTTTGGGACCTTGGCCAGCCAGCCGCGGACCCCTGACGACCTGGCC 5023
QY 1701 TAGTGTGGGGCGGCTTCGTGTACTGTCAAGACTGTGTGAGAGCTGTGAGCGCTCGCGTG 1760
Db 5022 TAGTGTGGGGCGGCTTCGTGTACTGTCAAGACTGTGTGAGAGCTGTGAGCGCTCGCGTG 4863
QY 1761 CTCAGCGGCGCAACCCCGCGCGGCTCTACTGTGAGCAGATGCCCTATCCGTGTAT 1820
Db 4962 CTCAGCGGCGCAACCCCGCGCGGCTCTACTGTGAGCAGATGCCCTATCCGTGTAT 4903
QY 1821 GTGGAAGACGTGTCTGTGCTGTGAGCGGCTGTGAGCGGCTGTGCTGTGAGCGTGGCC 1880
Db 4902 GTGGAAGACGTGTCTGTGCTGTGAGCGGCTGTGAGCGGCTGTGCTGTGAGCGTGGCC 4843
QY 1881 TGGATCTACTCCGTGACACTGACAGTGAAGGCGGTGTGTGGGAGAGGAGAGCGGCTG 1940
Db 4842 TGGATCTACTCCGTGACACTGACAGTGAAGGCGGTGTGTGGGAGAGGAGAGCGGCTG 4783
QY 1941 CGGGACACCATGGCGGCATGGGGCTCAGCGCGCGGTGTCTCTGGCTAGGCTGGTCCCTC 2000
Db 4782 CGGGACACCATGGCGGCATGGGGCTCAGCGCGCGGTGTCTCTGGCTAGGCTGGTCCCTC 4723
QY 2001 AGCTGCCCTGGGCGCTTCTGTCTGAGCGCGGTGTGTTCTGTGTCTCAAGCTGGGG 2060
Db 4722 AGCTGCCCTGGGCGCTTCTGTCTGAGCGCGGTGTGTTCTGTGTCTCAAGCTGGGG 4663
QY 2061 GACATCTCTCCCTACAGCACCCGGGGGTGTCTCTGTGTCTTGGCAGCTTCGGGCTG 2120
Db 4662 GACATCTCTCCCTACAGCACCCGGGGGTGTCTCTGTGTCTTGGCAGCTTCGGGCTG 4603
QY 2121 GCCACGCTGACCCAGAGCTTCTGTCTGAGCGCTTCTTCTCCGCGCCAACTGGGTGCG 2180
Db 4602 GCCACGCTGACCCAGAGCTTCTGTCTGAGCGCTTCTTCTCTCCGCGCCAACTGGGTGCG 4543

QY 2181 GCTGCGGGCGCTGSCCTACTTCTCCCTACTCTGCTGCTGCTGTGTGGCTTGG 2240
Db 4542 GCTGCGGGCGCTGSCCTACTTCTCCCTACTCTGCTGCTGCTGTGTGGCTTGG 4483
QY 2241 CGGGACCGGCTGCTCCCGGGGTGCGCGTGGCGCGAGCCTGTCTGCGCCGCTGGCC - TT 2299
Db 4482 CGGGACCGGCTGCTCCCGGGGTGCGCGTGGCGCGAGCCTGTCTGCTGCGCCGCTT 4423
QY 2300 CGGCTTTCGCTGG - AGAGCCTGGCTGTGTGAGAGAGAGGCGAGGCGCGCAGTGGC 2358
Db 4422 CGGCTTTCGCTGCAAGAGCCTGGCTCTGTGTGAGAGAGGCGAGGCGCGCAGTGGC 4363
QY 2359 ACNACTGGGACCCCGGCTACGGCAGAGCTTTCAGCCTGGCCAGGCTCTCTGGCCTTC 2418
Db 4362 ACACTGTGGACCCCGGCTTACGGCAGAGCTTTCAGCCTGGCCAGGCTCTCTGGCCTTC 4303
QY 2419 TGTGTGTGACCGCGGCTCTACGGCTCGCACCTGTGTACCTGGAAAGCTGTGTGCCAG 2478
Db 4302 TGTGTGTGACCGCGGCTCTACGGCTCGCACCTGTGTACCTGGAAAGCTGTGTGCCAG 4243
QY 2479 GGCAGTACGGGATTCCTGAACCAATGAAATTTTCTTTTCGGAGAGCTACTGTGTGGGAC 2538
Db 4242 GGCAGTACGGGATTCCTGAACCAATGAAATTTTCTTTTCGGAGAGCTACTGTGTGGGAC 4183
QY 2539 CTCGGCCCGGAGAGTCCAGCCCTTGGCCCGCCAGCTGGACCCAAAGGTCTGTGTAG 2598
Db 4182 CTCGGCCCGGAGAGTCCAGCCCTTGGCCCGCCAGCTGGACCCAAAGGTCTGTGTAG 4123
QY 2599 AAGAGCACCGCCCGGCTGAGTCTCTGGGCTATCCGTTTCGAGCCTGGAGAACGCTTTC 2658
Db 4122 AAGAGCACCGCCCGGCTGAGTCTCTGGGCTATCCGTTTCGAGCCTGGAGAACGCTTTC 4063
QY 2659 CTGGAAGCCGAGCCAGCCCTGCGGGGCTGAGCCTGTGACCTTCTTACAGGGCCACATCA 2718
Db 4062 CTGGAAGCCGAGCCAGCCCTGCGGGGCTCAGCCTGGACTTCTTACAGGGCCACATCA 4003
QY 2719 CGGCTTCTGCGGCACAAACGGGGCGGCAAGACCAACCCCTGTCTTCTTTCAGTGGCC 2778
Db 4002 CGGCTTCTGCGGCACAAACGGGGCGGCAAGACCAACCCCTGTCTTCTTTCAGTGGCC 3943
QY 2779 TCTTCCACACAGTGTGTGCTGTCTTTCATCTGCGGCCACGACGCTCCGCTCCAGCATGG 2838
Db 3942 TCTTCCACACAGTGTGTGCTGTCTTTCATCTTGGGCCACGACGCTCCGCTCCAGCATGG 3883
QY 2839 CGGCCATCCGCGCCCACTGTGGGCTGTCTCAGTACAACTGTGTGTGATGATGCTGA 2898
Db 3882 CGGCCATCCGCGCCCACTGTGGGCTGTCTCAGTACAACTGTGTGTGATGATGCTGA 3823
QY 2899 CGGTGACAGGACAGCTGTGTTCTATGGGGCTGAAAGGCTCTGAGTCCGCTGTAGTGG 2958
Db 3822 CGGTGACAGGACAGCTGTGTTCTATGGGGCTGAAAGGCTCTGAGTCCGCTGTAGTGG 3763
QY 2959 GCGCCAGCAGGACCTGTCTGCTGAGGATGTGGGGCTGTCTCCAAAGCAGAGTGTGAGA 3018
Db 3762 GCGCCAGCAGGACCTGTCTGCTGAGGATGTGGGGCTGTCTCCAAAGCAGAGTGTGAGA 3703
QY 3019 CTGCGCACTCTCTGTGTGGATGCAACGGAAGCTGTCCGTGGCCATGCTTTGTGGGGCG 3078
Db 3702 CTGCGCACTCTCTGTGTGGATGCAACGGAAGCTGTCCGTGGCCATGCTTTGTGGGGCG 3643
QY 3079 GCTCCCAAGTGTATCTCTGAGCGGCTACGGCTGCGGTGGATCTGTCTTCCGCGCGG 3138
Db 3642 KCKCCCAAGTGTATCTCTGAGCGGCTACGGCTGCGGTGGATCTGTCTTCCGCGCGG 3583
QY 3139 GTATTGGGAGCTGTGCTCAATAACGGAAGGTGCGACGCTGTATCTCTCCACCCACC 3198
Db 3582 GTATTGGGAGCTGTGCTCAATAACGGAAGGTGCGACGCTGTATCTCTCCACCCACC 3523
QY 3199 ACCTGATGAGGAGAGCTGTGCGGAGACCGTGTGCTGTGTGGAGAGTGGCGGCTTGT 3258
Db 3522 ACCTGATGAGGAGAGCTGTGCGGAGACCGTGTGCGGAGACCGTGTGGAGAGTGGCGGCTTGT 3463

QY 3259 GCTGCTGGCTCCCACTCTTCTGCGCGCTCACTGGGCTCCGGCTACTACTGACGC 3318
DB 3462 GCTGTGTGGCTCCCACTCTTCTGCGCGCTCACTGGGCTCCGGCTACTACTGACGC 3403
QY 3319 TGGTGAAGCCCGCTGCGCCCTGACCAACCAATGAGAAGCTGACACTGACATGGAAGGCA 3378
DB 3402 TGGTGAAGCCCGCTGCGCCCTGACCAACCAATGAGAAGCTGACACTGACATGGAAGGCA 3343
QY 3379 GTGTGGACACAGGAGGAAAGAAATGGCAGCCAGGCGACAGAGTCGSCACTCTCTC 3438
DB 3342 GTGTGGACACAGGAGGAAAGAAATGGCAGCCAGGCGACAGAGTCGSCACTCTCTC 3283
QY 3439 AGCTGCTGGCCCTGTGACAGCACTGGGTGCGCGGCGACGGCTGTGTGAGAGCTGCGAC 3498
DB 3282 AGCTGCTGGCCCTGTGACAGCACTGGGTGCGCGGCGACGGCTGTGTGAGAGCTGCGAC 3223
QY 3499 ACGAGCTGGTGTGTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCT 3558
DB 3222 ACGAGCTGGTGTGTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCT 3163
QY 3559 TCCGAGAGCTAGACACGGCGGTGGCGGAGCTGAGGCTCACTGGCTACCGGATCTCCGACA 3618
DB 3162 TCCGAGAGCTAGACACGGCGGTGGCGGAGCTGAGGCTCACTGGCTACCGGATCTCCGACA 3103
QY 3619 CCAGCTCGAGAGATCTTCTGAAAGTGTGTGAGAGTGTGCTGCGGACACAGATATGG 3678
DB 3102 CCAGCTCGAGAGATCTTCTGAAAGTGTGTGAGAGTGTGCTGCGGACACAGATATGG 3043
QY 3679 AGGATGGCAGCTGGGGCAGCACCTATGACAGGCAATTCCTGGCTAGACGTAACTCCCTGC 3738
DB 3042 AGGATGGCAGCTGGGGCAGCACCTATGACAGGCAATTCCTGGCTAGACGTAACTCCCTGC 2983
QY 3739 GSGCTCAAGATCCGACACAGGAGACAGCGCTGGAGAACGGGGAAACCACTGGGTGAGGCC 3798
DB 2982 GSGCTCAAGATCCGACACAGGAGACAGCGCTGGAGAACGGGGAAACCACTGGGTGAGGCC 2923
QY 3799 CAGAGACTGACAGGGCTCTGGGCGACAGCGCTGGGCGGGTACAGGGCTGGGCACTGA 3858
DB 2922 CAGAGACTGACAGGGCTCTGGGCGACAGCGCTGGGCGGGTACAGGGCTGGGCACTGA 2863
QY 3859 CCGCGCAGCAGCTCCAGGCGCTGCTTCTCAAGCGCTTCTGCTGGCGCGCAGCGGCC 3918
DB 2862 CCGCGCAGCAGCTCCAGGCGCTGCTTCTCAAGCGNGGGGGGGGGGGGGGGGG 2803
QY 3919 GCGGCGCTTTCGCGCCAGATCGTGTGCTGCTGCGCTCTTTGTGGGCGTGGCCCTCGTGTCA 3978
DB 2802 GCGGCGCTT----- 2793
QY 3979 GCCTCATGCTGCTCTTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCCACTGATAG 4038
DB 2792 ----- 2793
QY 4039 GTGCTCAGGTGCTTCTTTCAGTGAGGACGCCCAAGGGGACCTTGAGAGTCCCGGCTGC 4098
DB 2792 -----GTTTCGCCCACTGAGGACGCCCAAGGGGACCTTGAGAGTCCCGGCTGC 2745
QY 4099 TCGAGGCGCTGCTGACGAGGACGAGCTGAGGAGCGCCCACTGAGTCCCACTGAGTCCCA 4158
DB 2744 TCGAGGCGCTGCTGACGAGGACGAGCTGAGGAGCGCCCACTGAGTCCCACTGAGTCCCA 2685
QY 4159 GGTTCCTGGCACCAAGAGTTCCTGCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCACTGGA 4218
DB 2684 GGTTCCTGGCACCAAGAGTTCCTGCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCACTGGA 2625
QY 4219 CCCAGAGTCTCCATCCCACTGCTGCTAGCTAGCGAGCGCGGTGCGCGGCGCTGCTGC 4278
DB 2624 CCCAGAGTCTCCATCCCACTGCTGCTAGCTAGCGAGCGCGGTGCGCGGCGCTGCTGC 2565
QY 4279 CCGACTGCGCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4338
DB 2564 CCGACTGCGCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2505
QY 4339 AAGTGGTTTCAAGACCTGACAGGCGGAACTGTCTGACTTCTCTGCTCAAGACCTACCGC 4398

DB 2504 AAGTGGTTTCAAGACCTGACAGGCGGAACTGTCTGACTTCTGCTCAAGACCTACCGC 2445
QY 4399 GCCTGGTGCAGCGGCTGAAGACTAAGAAAGTGGTGAATCAGGTACAGGTACGAGGCT 4458
DB 2444 GCCTGGTGCAGCGGCTGAAGACTAAGAAAGTGGTGAATCAGGTACAGGTACGAGGCT 2385
QY 4459 TCTCGCTGGGGGCGGAGACCCAGGCTTGCCTCGGGCCAAAGAGTTGGGCGCTCAGTGG 4518
DB 2384 TCTCGCTGGGGGCGGAGACCCAGGCTTGCCTCGGGCCAAAGAGTTGGGCGCTCAGTGG 2325
QY 4519 AGGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCTGGGCGGCGCTGACCGTGTCTGAAAA 4578
DB 2324 AGGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCTGGGCGGCGCTGACCGTGTCTGAAAA 2265
QY 4579 ACCTCACAGCTGGCTCACAGCTTGAATGCTCAGGACAGTCTCAAGATCTGGTTCAACA 4638
DB 2264 ACCTCACAGCTGGCTCACAGCTTGAATGCTCAGGACAGTCTCAAGATCTGGTTCAACA 2205
QY 4639 ACAAAGGCTGGCACTTCCATGTGGCTTTTCAACCGAGCCAGCAACGCAATCTCTCCGTG 4698
DB 2204 ACAAAGGCTGGCACTTCCATGTGGCTTTTCAACCGAGCCAGCAACGCAATCTCTCCGTG 2145
QY 4699 CTCACTGCGCCCGAGGCGCGCGCCACAGCCACAGCATCACACACTCAACCACTCCCT 4758
DB 2144 CTCACTGCGCCCGAGGCGCGCGCCACAGCCACAGCATCACACACTCAACCACTCCCT 2085
QY 4759 TGAACCTCACCAAGGAGCAGCTGTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTG 4818
DB 2084 TGAACCTCACCAAGGAGCAGCTGTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTG 2025
QY 4819 TCGTCTCATCTGTGTGTGCTTTTGCCTGCTTGTCCCGGCGAGCTTCACTCTGTGTC 4878
DB 2024 TCGTCTCATCTGTGTGTGCTTTGCGCATGCTTGTGTCCCGGCGAGCTTCACTCTGTGTC 1965
QY 4879 TCATTGAGGAGCAGTCACTCCGAGCCAGCACTGACGCTCATGAGGCGGCTGTCCCCCA 4938
DB 1964 TCATTGAGGAGCAGTCACTCCGAGCCAGCACTGACGCTCATGAGGCGGCTGTCCCCCA 1905
QY 4939 CCCTCTACTGCTGGGCACTTCTCTGGGACATGCTAACTACTTGGTGGCAGCATGCA 4998
DB 1904 CCCTCTACTGCTGGGCACTTCTCTGGGACATGCTAACTACTTGGTGGCAGCATGCA 1845
QY 4999 TCGTGGTGTCTCATCTTCTGCGCTTCCAGCAGAGGCGATATGTGCGCCCTGCAACCTGC 5058
DB 1844 TCGTGGTGTCTCATCTTCTGCGCTTCCAGCAGAGGCGATATGTGCGCCCTGCAACCTGC 1785
QY 5059 CTGCTCTCTGCTGCTACTACTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 5118
DB 1784 CTGCTCTCTGCTGCTACTACTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1725
QY 5119 CCTCTCTTCTTCTGCTGCGGCGAGCAGGCTATGTGCTGCTCACTGCAATAAAGCTCT 5178
DB 1724 CCTCTCTTCTTCTGCTGCGGCGAGCAGGCTATGTGCTGCTCACTGCAATAAAGCTCT 1665
QY 5179 TTATTGCGCATCAATGGAAGCATGCGCACCTTTGTGCTTGTGCTTGTCTCTCATCAGAAGC 5238
DB 1664 TTATTGCGCATCAATGGAAGCATGCGCACCTTTGTGCTTGTGCTTGTCTCTCATCAGAAGC 1605
QY 5239 TGCAGGAGGTGAGCCGATCTTGAACAAGTCTTCTTATCTTCCCCACCTTCTGCTTGG 5298
DB 1604 TGCAGGAGGTGAGCCGATCTTGAACAAGTCTTCTTATCTTCCCCACCTTCTGCTTGG 1545
QY 5299 GCGGGGGCTCATTTGACATGCTGCGGAACAGGCGCATGGCTGATGCTTTGAGGCGCTTGG 5358
DB 1544 GCGGGGGCTCATTTGACATGCTGCGGAACAGGCGCATGGCTGATGCTTTGAGGCGCTTGG 1485
QY 5359 GAGACGCGAGTTCCAGTCACTCCCTGCTGGGAGGTGGTGGCAAGAACCTCTTTGGCCA 5418
DB 1484 GAGACGCGAGTTCCAGTCACTCCCTGCTGGGAGGTGGTGGCAAGAACCTCTTTGGCCA 1425
QY 5419 TGGTGTATACAGGGGCGCTCTTCTTCTTCACTACTGCTGACGACCGAGGCGCAAC 5478

QY 4410 CAGGSCCTGAAGACTAAGAGTGGGTGAATGAGGTTCAGGTACGAGAGGCTTCTCGTGGGG 4469
Db |||||
QY 4420 CAGGSCCTGAAGACTAAGAGTGGGTGAATGAGGTTCAGGTACGAGAGGCTTCTCGTGGGG 4260
Db |||||
QY 4470 GCGCGAGACCCAGAGGCTCGCCCTCGGGCCCAAGAGTTGGGCGCCTCAGTCGAGGAGCTTGTGG 4529
Db |||||
QY 4261 GCGCGAGACCCAGAGGCTCGCCCTCGGGCCCAAGAGTTGGGCGCCTCAGTCGAGGAGCTTGTGG 4320
Db |||||
QY 4530 GCGCTGCTGAGTCCCTCGCTCGGGGGGCCCTCGACCGTGTCTGAAACAACTCACAGCC 4589
Db |||||
QY 4321 GCGCTGCTGAGTCCCTCGCTCGGGGGGCCCTCGACCGTGTCTGAAACAACTCACAGCC 4380
Db |||||
QY 4590 TGGGCTCAAGCTGGATGCTCAGGACAGTCTCAAGATCTGGTTTCAACAAAGGCTGG 4649
Db |||||
QY 4381 TGGGCTCAAGCTGGATGCTCAGGACAGTCTCAAGATCTGGTTTCAACAAAGGCTGG 4440
Db |||||
QY 4650 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCCTCGTGCTCACTGCC 4709
Db |||||
QY 4441 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCCTCGTGCTCACTGCC 4500
Db |||||
QY 4710 CCAGGCCCGGCCCGCCAGCCACAGCATCACCACTCAACACCCCTTGAACTCAC 4769
Db |||||
QY 4501 CCAGGCCCGGCCCGCCAGCCACAGCATCACCACTCAACACCCCTTGAACTCAC 4560
Db |||||
QY 4770 AAGGAGCAGCTGCTGAGGCTGCACTGATGCGCTCTCGGTGGACGCTCTCGTCCATC 4829
Db |||||
QY 4561 AAGGAGCAGCTGCTGAGGCTGCACTGATGCGCTCTCGGTGGACGCTCTCGTCCATC 4620
Db |||||
QY 4830 TGTGTGCTTTGGCATGCTCTTGTCCCGGCCAGCTTCACTTGTCTTCAATGAGGAG 4889
Db |||||
QY 4621 TGTGTGCTTTGGCATGCTCTTGTCCCGGCCAGCTTCACTTGTCTTCAATGAGGAG 4680
Db |||||
QY 4890 CGAGTCAACCCAGCAAGCACTGCGAGCTCATGGGGGCTGTCCCCACCCCTTACTGG 4949
Db |||||
QY 4681 CGAGTCAACCCAGCAAGCACTGCGAGCTCATGGGGGCTGTCCCCACCCCTTACTGG 4740
Db |||||
QY 4950 CTTGGCAACTTCTCTGGGCAATGTGTAATCTCTGTTGGTCCAGCATGCTGGTGTCTC 5009
Db |||||
QY 5010 ATCTTTCTGGCTTCCAGCAGAGGCGCATATGTGGCCCTGCAACCTGCTCTCTCTG 5069
Db |||||
QY 4801 ATCTTTCTGGCTTCCAGCAGAGGCGCATATGTGGCCCTGCAACCTGCTCTCTCTG 4860
Db |||||
QY 5070 CTGTTGCTACTGATGATGCTGCTCGATCACCGCTCATGTACCCAGCCTCTCTCTC 5129
Db |||||
QY 4861 CTGTTGCTACTGATGATGCTGCTCGATCACCGCTCATGTACCCAGCCTCTCTCTC 4920
Db |||||
QY 5130 TTCTCCGTGCCAGCAGCCATATGTGCTGCTACCTGCAATAAACCTCTTTATTGGCATC 5189
Db |||||
QY 4921 TTCTCCGTGCCAGCAGCCATATGTGCTGCTACCTGCAATAAACCTCTTTATTGGCATC 4980
Db |||||
QY 5190 AATGGAAGCATGGCCACCTTTGCTGCTGCTCTCTGATCAGAGCTGCAGAGGCTG 5249
Db |||||
QY 4981 AATGGAAGCATGGCCACCTTTGCTGCTGCTCTCTGATCAGAGCTGCAGAGGCTG 5040
Db |||||
QY 5250 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGCACTTCTGCTGGGCGGGGCTC 5309
Db |||||
QY 5041 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGCACTTCTGCTGGGCGGGGCTC 5100
Db |||||
QY 5310 ATTGACATGCTGGGAAACCAAGGCCATGGCTGATCGCTTTTGAAGCGCTTTGGGAGACAGGCG 5369
Db |||||
QY 5101 ATTGACATGCTGGGAAACCAAGGCCATGGCTGATCGCTTTTGAAGCGCTTTGGGAGACAGGCG 5160
Db |||||
QY 5370 TTCCAGTCAACCCCTGCGCTGGAGGCTGCTGGCAAGAACCTTTGGCGCATGGTATACAG 5429
Db |||||
QY 5161 TTCCAGTCAACCCCTGCGCTGGAGGCTGCTGGCAAGAACCTTTGGCGCATGGTATACAG 5220
Db |||||
QY 5430 GGGCCCTCTTCTCTTCTTCCACTACTGCTGTCGAGCAGCCGAGCAACCTCTGCCACAG 5489
Db |||||
QY 5221 GGGCCCTCTTCTCTTCTTCCACTACTGCTGTCGAGCAGCCGAGCAACCTCTGCCACAG 5280
Db |||||

QY 5490 CCCAGGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAACGG 5549
Db |||||
QY 5281 CCCAGGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAACGG 5340
Db |||||
QY 5550 GAGCGGGTGGTCCCAAGGAGCCACCCAGGGGATGTGTTGGTGTCTGAGGAACTTGACCAAG 5609
Db |||||
QY 5341 GAGCGGGTGGTCCCAAGGAGCCACCCAGGGGATGTGTTGGTGTCTGAGGAACTTGACCAAG 5400
Db |||||
QY 5610 GTATACCGTGGGAGAGGATGCCAGCTGTTGACCGCTTGTGCTGGGATTTCCCCCTGGT 5669
Db |||||
QY 5401 GTATACCGTGGGAGAGGATGCCAGCTGTTGACCGCTTGTGCTGGGATTTCCCCCTGGT 5460
Db |||||
QY 5670 GAG 5672
Db |||||
QY 5461 GAG 5463
Db |||||
RESULT 15
ABX95285
ID ABX95285 standard; cDNA; 5859 BP.
AC ABX95285;
XX
XX AC
XX
DT 18-JUN-2003 (first entry)
XX
DE Human ABCA7 associated cDNA sequence #1.
XX
DE Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic; ss.
XX
OS Homo sapiens.
XX
XX
PN WO2003010315-A1.
XX
PD 06-FEB-2003.
XX
PF 24-JUL-2002; 2002WO-JP007487.
XX
PR 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX
PA (KYOW) KYOMA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX
DR WPI; 2003-239444/23.
XX
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
XX
PS Example 1; Page 154-158; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents a human cDNA sequence given in the
CC specification of the present invention
XX
SQ Sequence 5859 BP; 987 A; 1891 C; 1827 G; 1154 T; 0 U; 0 Other;
Query Match 71.2%; Score 4843.4; DB 8; Length 5859;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4911; Conservative 0; Mismatches 6; Indels 54; Gaps 1;
QY 702 CGCAGCGAATCCCTGGGGTTGGCACTGGGGCAAGCCAGGAGCCCTTTCACAGCTTGTG 761

Db 751 CCCAGGAATCCCTGGGGTGGCACTGGGCAAGCCAGAGGCCCTTGACACGTTGTG 810
Qy 762 GAGGCGCTGAGGACCTGGGCCCAAGAGCTCCTGGCGTGGCAGCCTGTGTGAGCTTGG 821
Db 811 GAGGCGCTGAGGACCTGGGCCCAAGAGCTCCTGGCGTGGCAGCCTGTGTGAGCTTGG 870
Qy 822 GCACTGCTGAGAGACCCGAGAGGACGAGGCGCCCTGGAGTTGTGTGAGAGGCCCTC 881
Db 871 GCACTGCTGAGAGACCCGAGAGGACGAGGCGCCCTGGAGTTGTGTGAGAGGCCCTC 930
Qy 882 TGCACTGTTCAGGGGACCTAGCAGCACAGTGGGCGCCCTCCCTCAACTGGTACAGAGCTAGT 941
Db 931 TGCACTGTTCAGGGGACCTAGCAGCACAGTGGGCGCCCTCCCTCAACTGGTACAGAGCTAGT 990
Qy 942 GACCTGATGAGTGTGTGGGCGAGGACCAAGATCCGCGCTGCCAGACAGCAGCCTGAGC 1001
Db 991 GACCTGATGAGTGTGTGGGCGAGGACCAAGATCCGCGCTGCCAGACAGCAGCCTGAGC 1050
Qy 1002 CCGCGCTGCTCGAGCTGATTGAGAGCCCTGGACAGCACCCGCTGTCGCGCTGCTGG 1061
Db 1051 CCGCGCTGCTCGAGCTGATTGAGAGCCCTGGACAGCACCCGCTGTCGCGCTGCTGG 1110
Qy 1062 AGACGCTCAAGGCTCTGTATCTCGGGAAGCTACTCTTTGCAACGATACACCTTTTACC 1121
Db 1111 AGACGCTCAAGGCTCTGTATCTCGGGAAGCTACTCTTTGCAACGATACACCTTTTACC 1170
Qy 1122 CGAAGCTCATGCGCCAGGTGAACCGGACCTTCGAGAGTCACTCCTTCATGAACGACGTTCCNAT 1181
Db 1171 CGAAGCTCATGCGCCAGGTGAACCGGACCTTCGAGAGTCACTCCTTCATGAACGACGTTCCNAT 1230
Qy 1182 CGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAACGACGTTCCNAT 1241
Db 1231 CGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAACGACGTTCCNAT 1290
Qy 1242 GTGGCCATGCTGACGGGCTCCTGAGATGACGAGTGAAGGAAGAGGACGCCAGACCT 1301
Db 1291 GTGGCCATGCTGACGGGCTCCTGAGATGACGAGTGAAGGAAGAGGACGCCAGACCT 1350
Qy 1302 GGAGGCGGAGACCAATGAGGCGCTGCGATCTTTCTGAGACCTCGGAGCGTGTGCTAC 1361
Db 1351 GGAGGCGGAGACCAATGAGGCGCTGCGATCTTTCTGAGACCTCGGAGCGTGTGCTAC 1410
Qy 1362 AGCTGGCAGACGACACGCTGATGTGGGCACTTGTGGGCACTGTGGGCGAGTGAAG 1421
Db 1411 AGCTGGCAGACGACACGCTGATGTGGGCACTTGTGGGCACTGTGGGCGAGTGAAG 1470
Qy 1422 GAGTGTCTGCTTGGACAAGCTGGAGGCGGCACTTCAAGGACGCTTGTGTGCGGG 1481
Db 1471 GAGTGTCTGCTTGGACAAGCTGGAGGCGGCACTTCAAGGACGCTTGTGTGCGGG 1530
Qy 1482 GCCCTGCACTGCTCGCGGAACATCGATTCGCGCGGCGTCTGCTTCTTGGACCTGAG 1541
Db 1531 GCCCTGCACTGCTCGCGGAACATCGATTCGCGCGGCGTCTGCTTCTTGGACCTGAG 1590
Qy 1542 GACTTTCAGACCCGACAGACCCCAACCCGAGACCTGGGCGCCGCGCACTGCGCATC 1601
Db 1591 GACTTTCAGACCCGACAGACCCCAACCCGAGACCTGGGCGCCGCGCACTGCGCATC 1650
Qy 1602 AAAATCCGATGACATTTGACGTGGTCAAGGACCAATGAAGATCAGGGACAGGTTTGG 1661
Db 1651 AAAATCCGATGACATTTGACGTGGTCAAGGACCAATGAAGATCAGGGACAGGTTTGG 1710
Qy 1662 GACCTGGCCGACCGCGGACCCCTGACCGACTCGGCTGCTCAGCGCGCCACCCCGG 1721
Db 1711 GACCTGGCCGACCGCGGACCCCTGACCGACTCGGCTGCTCAGCGCGCCACCCCGG 1770
Qy 1722 TACCTGCAAGACTGTGGAGCGTGCAGCGCTCGGCTGCTCAGCGCGCCACCCCGG 1781
Db 1771 TACCTGCAAGACTGTGGAGCGTGCAGCGCTCGGCTGCTCAGCGCGCCACCCCGG 1830
Qy 1782 GCCGCGCTCTACCTGACGAGATGCCCTATCCGCTGCTATGTGACGACGTTGTTCTGCGT 1841

Db 1831 GCCGCGCTCTACCTGACGAGATGCCCTATCCGCTGCTATGTGACGACGTTTCTGCGT 1890
Qy 1842 GTGCTGAGCCGCTGCTGCGGCTCTTCTCTGACGCTGGCTGGATCTACTCCGTTGACACTG 1901
Db 1891 GTGCTGAGCCGCTGCTGCGGCTCTTCTCTGACGCTGGCTGGATCTACTCCGTTGACACTG 1950
Qy 1902 ACAGTGAAGCGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1961
Db 1951 ACAGTGAAGCGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2010
Qy 1962 GGGCTCAGCGCGCGCTGCTCTGGTAGGCTGGTTCTCCTCAGCTGCTCGGCGCCCTTCTG 2021
Db 2011 GGGCTCAGCGCGCGCTGCTCTGGTAGGCTGGTTCTCCTCAGCTGCTCGGCGCCCTTCTG 2070
Qy 2022 CTCAGCGCGCGCTGCTGCTTCTGGTAGGCTGGTTCTCCTCAGCTGCTCGGCGCCCTTCTG 2081
Db 2071 CTCAGCGCGCGCTGCTGCTTCTGGTAGGCTGGTTCTCCTCAGCTGCTCGGCGCCCTTCTG 2130
Qy 2082 CCGGGCGTGGTCTTCTTCTTGGCAGGCTTTCGCGGTGGCCACGCTGACCCAGAGCTTC 2141
Db 2131 CCGGGCGTGGTCTTCTTCTTGGCAGGCTTTCGCGGTGGCCACGCTGACCCAGAGCTTC 2190
Qy 2142 CTGCTCAGCGCGCTTCTTCTTCCGCGCAACCTGGCTGCGGCTGCGGCGGCTGCGGCGTAC 2201
Db 2191 CTGCTCAGCGCGCTTCTTCTTCCGCGCAACCTGGCTGCGGCTGCGGCGGCTGCGGCTTAC 2250
Qy 2202 TTCTCCCTCTACCTGCGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2261
Db 2251 TTCTCCCTCTACCTGCGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2310
Qy 2262 GGGCGGTGGCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2321
Db 2311 GGGCGGTGGCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2370
Qy 2322 GCTCTGCTGGAGAGCAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACG 2381
Db 2371 GCTCTGCTGGAGAGCAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACG 2430
Qy 2382 GAGAGCTCTTACGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAC 2441
Db 2431 GAGAGCTCTTACGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAC 2490
Qy 2442 GGGCTCGCACCTGCTGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2501
Db 2491 GGGCTCGCACCTGCTGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2550
Qy 2502 TGGAAATTTTCTTCTTCTCGAGGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2561
Db 2551 TGGAAATTTTCTTCTTCTCGAGGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2610
Qy 2562 CTTTGGCCCAACCCGCTGAGACCCAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2621
Db 2611 CTTTGGCCCAACCCGCTGAGACCCAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2670
Qy 2622 CTTTGGGCTATCCGTTTCGAG 2681
Db 2671 CTTTGGGCTATCCGTTTCGAG 2730
Qy 2682 CCGGGGCTCAGGCTGAGCTTCTACAGGGGCCAATCACCGCCCTTCTGCGGCGCAACAGGG 2741
Db 2731 CCGGGGCTCAGGCTGAGCTTCTACAGGGGCCAATCACCGCCCTTCTGCGGCGCAACAGGG 2790
Qy 2742 GCGGCAAGACCAACCTGCTTCTTGGAGTGGCTCTTCTCCACCCAGTGGTGGCTCT 2801
Db 2791 GCGGCAAGACCAACCTGCTTCTTGGAGTGGCTCTTCTCCACCCAGTGGTGGCTCT 2850
Qy 2802 GCCTTCATCTGGGCGCACGAGCTCCGCTCCAGATGGCGGCCATCCGCGGCCCACTCTGGGC 2861
Db 2851 GCCTTCATCTGGGCGCACGAGCTCCGCTCCAGATGGCGGCCATCCGCGGCCCACTCTGGGC 2910
Qy 2862 GTCTGTCTCAGTACAACTGCTGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2921
Db 2911 GTCTGTCTCAGTACAACTGCTGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2970

Db	5131	 CTGTATG-----	5137
Qy	5142	AGCAGCGCTATGTGTGCTCACCTGCATAAACCTCTTTATTGGCATCAATGGAAGCATG	5201
Db	5138	-GCACGCCATGTGGTGCTCACCTGCATAAACCTCTTTATTGGCATCAATGGAAGCATG	5196
Qy	5202	GCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGTGAGCCCGATCTTG	5261
Db	5197	GCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGNAGCTGCAGAGGTGAGCCCGATCTTG	5256
Qy	5262	AAACAGGCTTTCCTTATCTTCTCCCACTCTCTGCTGGGCCGGGGCTCATTTGACATGGTG	5321
Db	5257	AAACAGGCTTTCCTTATCTTCTCCCACTCTCTGCTGGGCCGGGGCTCATTTGACATGGTG	5316
Qy	5322	CGGAACAGGCGCATGGCTGTATGCTTTGAGCGCTTGGGAGACAGCGATGTCAGTCAACCC	5381
Db	5317	CGGAACAGGCGCATGGCTGTATGCTTTGAGCGCTTGGGAGACAGCGATGTCAGTCAACCC	5376
Qy	5382	CTGGCGCTGGGAGGTGCTCGGCAAGAACTCTTGGCCATGTGTATACAGGGGCCCCCTTTC	5441
Db	5377	CTGGCGCTGGGAGGTGCTCGGCAAGAACTCTTGGCCATGTGTATACAGGGGCCCCCTTTC	5436
Qy	5442	CTTCTCTTTACACTACTGCTGCACGCCAGGCCAACTCTTCGCCACAGCCAGGGTGAAG	5501
Db	5437	CTTCTCTTTACACTACTGCTGCACGCCAGGCCAACTCTTCGCCACAGCCAGGGTGAAG	5496
Qy	5502	TCTCTGCCACTCTCTGGAGAGGAGACGAGGATGTAGCCCGTGAAACGGGAGCGGGTGGTC	5561
Db	5497	TCTCTGCCACTCTCTGGAGAGGAGACGAGGATGTAGCCCGTGAAACGGGAGCGGGTGGTC	5556
Qy	5562	CAAGGAGCCACCCAGGGGGATGTTGGTGCTGAGGAACTTGACCAAGGTATACCGTGGG	5621
Db	5557	CAAGGAGCCACCCAGGGGGATGTTGGTGCTGAGGAACTTGACCAAGGTATACCGTGGG	5616
Qy	5622	CAGAGGATGCCAGCTGTTGACCGCTTGTGCTTGGGGATTCCCCCTGGTGAG	5672
Db	5617	CAGAGGATGCCAGCTGTTGACCGCTTGTGCTTGGGGATTCCCCCTGGTGAG	5667

Search completed: December 29, 2004, 23:46:36
Job time : 2933 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:51:51 ; Search time 493 Seconds
(without alignments)
9809.752 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctcaggggcgcgctcccc.....cgtgctggtgaaaaa 6804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737.4	25.5	7860	US-09-526-193A-2	Sequence 2, Appli
2	1368	20.1	6819	US-09-032-438C-2	Sequence 2, Appli
3	1368	20.1	7488	US-09-032-438C-1	Sequence 1, Appli
4	1291.4	19.0	6705	US-09-032-438C-5	Sequence 5, Appli
5	920	13.5	8037	US-09-774-528-209	Sequence 209, App
6	404.6	5.9	5894	US-08-665-259-24	Sequence 24, Appl
7	404.6	5.9	5894	US-08-762-500-24	Sequence 24, Appl
8	404.6	5.9	5894	US-08-762-500-74	Sequence 24, Appl
9	161.2	2.4	987	US-09-724-797-7	Sequence 74, Appl
10	158.6	2.3	1062	US-10-000-489-101	Sequence 101, App
11	143.4	2.1	447	US-09-621-976-13889	Sequence 13889, A
12	134.2	2.0	4403765	US-09-103-840A-2	Sequence 2, Appli
13	134.2	2.0	4411529	US-09-103-840A-1	Sequence 1, Appli
14	125.2	1.8	999	US-09-252-991A-4588	Sequence 4588, Ap
15	125.2	1.8	1008	US-09-252-991A-4266	Sequence 4266, Ap
16	120	1.8	36181	US-08-311-731A-120	Sequence 120, App
17	119.8	1.8	20284	US-09-526-193A-21	Sequence 21, Appl
18	119.2	1.8	1767	US-09-489-039A-6408	Sequence 6408, Ap
19	118.4	1.7	1614	US-09-252-991A-6293	Sequence 6293, Ap
20	118.4	1.7	2790	US-09-252-991A-6375	Sequence 6375, Ap
21	118.4	1.7	3948	US-09-252-991A-6175	Sequence 6175, Ap
22	116	1.7	1824	US-09-252-991A-1780	Sequence 1780, Ap
23	113.8	1.7	762	US-09-252-991A-1938	Sequence 1938, Ap
24	110.8	1.6	1614	US-09-252-991A-4198	Sequence 4198, Ap
25	110.2	1.6	4244	US-09-526-193A-27	Sequence 27, Appl
26	107.4	1.6	930	US-09-252-991A-4633	Sequence 4633, Ap
27	107.4	1.6	1179	US-09-252-991A-4241	Sequence 4241, Ap

28	107.4	1.6	1719	4	US-09-252-991A-4534	Sequence 4534, Ap
29	102.8	1.5	2769	4	US-09-489-039A-1767	Sequence 1767, Ap
30	101.2	1.5	6289	4	US-09-526-193A-26	Sequence 26, Appl
31	101.2	1.5	536165	4	US-09-214-808-1	Sequence 1, Appli
32	92	1.4	1176	4	US-09-252-991A-3704	Sequence 3704, Ap
33	92	1.4	2253	4	US-09-252-991A-3810	Sequence 3810, Ap
34	91	1.3	1110	4	US-09-252-991A-8962	Sequence 8962, Ap
35	91	1.3	1278	4	US-09-252-991A-9043	Sequence 9043, Ap
36	91	1.3	1968	4	US-09-252-991A-8743	Sequence 8743, Ap
37	90.4	1.3	750	4	US-09-252-991A-795	Sequence 795, App
38	90.4	1.3	786	4	US-09-252-991A-872	Sequence 872, App
39	90.4	1.3	1197	4	US-09-252-991A-808	Sequence 808, App
40	87.2	1.3	972	4	US-09-489-039A-5325	Sequence 5325, Ap
41	87.2	1.3	1047	4	US-09-602-787A-347	Sequence 347, App
42	86.6	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
43	86.6	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
44	85.8	1.3	813	4	US-09-602-787A-349	Sequence 349, App
45	84.6	1.2	35881	4	US-08-311-731A-127	Sequence 127, App

ALIGNMENTS

RESULT 1
US-09-526-193A-2
; Sequence 2, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pimstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526.193A
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-526-193A-2

Query Match	25.5%	Score 1737.4;	DB 4;	Length 7860;
Best Local Similarity	58.8%	Pred. No. 0;		
Matches 3342;	Conservative	0;	Mismatches 2191;	Indels 150; Gaps 14;
Qy	987	GACAGCAGCTGACCGCGCTGCTCGAGCTGATTGGAGCCCTGGGACGACCCCGCTG	1046	
Db	1116	GACACTCTACACTCTTACTGCAATGATTGTAAGAAATTTGGAGCTAGTCTCTT	1175	
Qy	1047	TCCGCGCTGCTCTGGAGACGCTGAGCCTCTGATCTCTGGGAAAGTACTCTTTGCACCA	1106	
Db	1176	TCCGCGATTATCTGGAAGCTCTGAGCGCTGCTCGTGGAGAGATCTGTATACACT	1235	
Qy	1107	GATACACTTTTATCCCGGAGAGCTCATGCGCCAGGTAAACCGGACCTTCGAGAGCTCACC	1166	
Db	1236	GACACTCCAGCCACAGGCGAGGTATGGCTGAGGTAAACAAGACCTTCCAGGAACCTGCT	1295	
Qy	1167	CTGCTGAGGATGTCGCGGAGGTGTGGAGATGCTGGAGCCCGGATCTTCCACCTTCATG	1226	
Db	1296	GTGTTCCATGATCTGGAAGGATGTGGAGAACTAGCCCCCAAGATCTGGACCTTCATG	1355	
Qy	1227	AACGACAGTTTCAATGTGGCATGCTGCAGCGGCTCTCTGCAGATGCAGGA-----	1276	


```
Db 5712 CCTGTAATCAAAGCTATCTCTCTGAATGATGAAGATGTGAGCGGGAAGA 5771
Qy 5550 GAGCGGGTGTCCAAGAGAGCCACCCAGGGGATGTGTGTGTGTGAGGAATTTGACCAAG 5609
Db 5772 CAGAGAATTTCTGTATGTGGAGGCCAGATGACATCTTAGAATCAAGAGATTTGACGAAG 5831
Qy 5610 GTATACCGTGGCAGAGATGCCAGCTGTGTGACCGCTTGTGCTGGGATTTCCCTCGGT 5669
Db 5832 ATATATAGAAGGAAGCGGAAGCCCTGTGTGACAGGATTTGCGTGGGCATTTCTCTCGGT 5891
Qy 5670 GAGTGTTTGGGCTGCTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTTCATGGTG 5729
Db 5892 GAGTGTTTGGGCTCTGGAGATTAATGGGCTGGAATATCAATCAATTTCAAGATGTTA 5951
Qy 5730 ACGGGGACACATTTGGCCACAGAGGGCGAGGCTGTGCTGCGAGGCCACAGCGTGGCCGG 5789
Db 5952 ACAGGAGATACCACTGTTTACAGAGGAGATGCTTTCTTAAACAAAATAGTATCTTATCA 6011
Qy 5790 GAACCCAGTCTGGCACCTCAGCATGGGATCTGCCCTCAATCCGATGCACTTTTGAG 5849
Db 6012 AACATCATGAAGTACATCAACATGGGCTACTGCCCTCAGTTTGATGCCATCACAGAG 6071
Qy 5850 CTGCTGACGGCGCGAGCAGCACTGGAGCTGTTGCGGCGCTGCGCGGTGTCCCGAGGCC 5909
Db 6072 CTGTTGACTGGAGAGACACGTGGAGTTCTTTGGCCCTTTTGAGAGGATCCCAGAGAA 6131
Qy 5910 CAGTTGCCAGACCGCTGGCTCAGGCCCTGCGCGGTCTGGGACTCTCANGTACGCAGAC 5969
Db 6132 GAAAGTTGGCAAGTTGCTGAGTGGCGATTTCGGAACCTGGGCTCGTGAAGTATGGAGAA 6191
Qy 5970 CGGCTGCAGGCACCTACAGCGAGGGAACAAACGCAAGCTGGGAGCGCCCTGGCGCTG 6029
Db 6192 AAATATGCTGTTAACTATAGTAGGAGCAACAAACGCAAGCTCTTACAGCCATGGCTTG 6251
Qy 6030 GTTGGGACCCAGCGGTGTTTCTGAGCAGCGACCAACAGCATGGACCCCGAGCGG 6089
Db 6252 ATCGCGGGCTCTGTGTGTGTTCTGATGAAACCCACACAGCATGGATGCCAAGCC 6311
Qy 6090 CGCGCTCTCTTTGGAAACAGCCTTTGGCCGTGTGGCGGAGGCGGTTCAGTGTGCTC 6149
Db 6312 CGCGGGTCTTGTGGAATTTGGCCCTAAGTGTGTCAGAGGGGAGATCAGTAGTCTT 6371
Qy 6150 ACCTCCATGATGAGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGGTGAAT 6209
Db 6372 ACATCTCATAGTATGGAAGATGTGAAGCTTTTGCACTAGGATGGCAATCATGGTCAAT 6431
Qy 6210 GGGCGGTTCCGCTGCGCAGCGCCCAACATCTCAAGGGCAGATTCGCGGCGGTCCAC 6269
Db 6432 GGAAGTTTCAAGTGTGCTTGGCAGGTGTCAGCATCTAATAAATAGTTTGGAGATGTTAT 6491
Qy 6270 AACTGACCTTCGG- - - - -GGTGCCCGCCGCAAGTTCACGCGGCGAGCGGCTTCGTG 6323
Db 6492 ACAATAGTTGTACGAATAGCAGGCTCAACCCGAGCCTGAAGCCTGTCCAGGATTTCTTT 6551
Qy 6324 CGGCGGAGTTCCTGTGGTGGAGCTCGGAGGCACATGGAGGCCGCTGCGCTTCAG 6383
Db 6552 GGACTTGCAATTTCTGGAAGTGTCTAANAAGAGAAACACCGGAACATGCTACAAATACCAG 6611
Qy 6384 CTGCGCGCGGAGGGCGCTGCGCCCTGCGCGCTCTTTTGGAGAGCTGGCGGTGCACGGC 6443
Db 6612 CTTCCATCTTCA- - - - -TTATCTTCTGCGCAGGATTTAGCATCTCTCCAGAGCAAA 6568
Qy 6444 GCAGAGACCGCGTGGAGGACTTTTCCGTGAGCGACGATGCTGGAGGAGTATCTTG 6503
Db 6669 AAGGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACTTGACCAAGTATTGTG 6728
Qy 6504 TACTTCTCCAGGACCAAGGAGGAGGACGAGACACCCGAAGC 6546
Db 6729 AACTTTGCCAAGGACCAAGTGTATGATGACCACTTAAAGACC 6771
```

```
US-09-032-438C-2
; Sequence 2, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 6819
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-032-438C-2
```

Query Match 20.1%; Score 1368; DB 4; Length 6819;
Best Local Similarity 54.4%; Pred. No. 8.9e-277;
Matches 3158; Conservative 0; Mismatches 2505; Indels 147; Gaps 14;

```
Qy 859 TGGAGTTGCTGTGCAGAGGCCCTCTGCAGTGTCTAGGGGACCTTAGCAGCAGACATGGGCCCT 918
Db 947 TGGGCACTCTGTCTGACCTCTGTGTGGCTACCCCGAGGAGGTGCTCTCGGGTGTCT 1006
Qy 919 CCCTCAACTGGTACGAGGCTTAGTGACCTGATGGAGCTGTGGGCGAGGAGCCAGAATCCG 978
Db 1007 CCTTCAACTGGTATGAAGCAATAACTATAGGCCCTTTCTGGGGATTGACTCCACAAGGA 1066
Qy 979 CCTGTC- - - - -AGACGAGCCTGAGCCCGCTCTCGGAGCTGATGGAG 1026
Db 1067 AGGATCCTATCTATTCTTATGACAGAAGAACACATCTTTTGTAAATGCAATTGATCCAGA 1126
Qy 1027 CCTGACAGCCACCGCTGTCCGCTGCTCTGGAGACGCTTGAAGCCTCTGATCCTCG 1086
Db 1127 GCCTGAGTCAAAATCTTTAAACCAAAATCGTTGGAGGGCGGCAAGCCTTTGCTGATGG 1186
Qy 1087 GGAAGCTACTCTTTTGACCAAGATACACCTTTTACCGGGAAGCTCATGGCCAGGTGAACC 1146
Db 1187 GAAAATCTGTACACTCTCTGATTCTACCTGCACACGAGGATCTGAAGATGCCAAT 1246
Qy 1147 GGAACCTTGGAGAGCTCACCTCTGAGGATGTCCGGAGGTGTGGGAGATGCTGGGAC 1206
Db 1247 CAACCTTTGAAGAACTGGAAACAGTTAGGAAGTTGGTCAAGCCTGGGAAGAAAGTAGGGC 1306
Qy 1207 CCGGATCTTCACTTTCATGACGAGTTCCAAATGTGGCCATGCTGCGAGCGCTCTGCG 1266
Db 1307 CCAGATCTGTGATCTTCTTTGACACAGCAACAGATGAACATGATCAGAGATACCTCGG 1366
Qy 1267 AGATGAGGATGAAGGAAGAGCCAGCAGCTGGAGGCCGAGCAAT- - - - -GG 1320
Db 1367 GGAACCCACAGTAAAGACTTTTGAATAGGCAGCTTGTGTGAAGAGGTATTACTGCTG 1426
Qy 1321 AGGCCCTGCGATCCTTTCTTGACCCCTGGGAGCGGTGCTACAGCTGGCAGGACGACAG 1380
Db 1427 AAGCCATCTTAAACTTCTCTACAAGGGCCCTCGGAAAGCCAGGCTGACGACATGATGCGCA 1486
```

QY 1381 CTGATGTGGGGCACCTGGTGGGACGCTGGGGCGAGTGACGAGTGCCTGTCTTGGACA 1440
DB 1487 ACTTCGACTGGAGGACATATTTAAACATCACTGATCGCACCTCCGCTGGTCAATCAAT 1546
QY 1441 AGCTGAGCGCGCACCTCCAGAGGAGCCCTGGTGTGGGGCCCTGCAACCTGCTCGGG 1500
DB 1547 ACTTGAGTGTGTCTGAGTAAAGTTTGAAGCTACAAATGATGAACCTCAGCTCACCC 1606
QY 1501 AACATCGATTCTGGGCGGCGTGTCTTCTTGGGACCTGAGGACTCTTCAGACCCACAG 1560
DB 1607 AAGTGGCCCTCTCTACTCTGAGGAAACATGTTCTGGGCGGAGTGGTATTCCTTGACA 1666
QY 1561 AGCACCCAAACCCAGACCTGGGCGCGGCGCACGCTGGGCAACAAATCCGCATGAGCATTTG 1620
DB 1667 TGTATCTCCCTGGACCACTCTCTACCAACCCACGTTGAAGTATGAAGATGGACATAG 1726
QY 1621 AGCTGGTCACGAGGACCAATAGATCAGGACAGGTTTGGGACCTGCGCCAGCGCGG 1680
DB 1727 AGCTGGTGAGAAACCAATAGATTAAAGACAGGTATTTGGGATTTCTGTCCAGAGCTG 1786
QY 1681 ACCCCCTGACCGACCTGCGCTACGTGTGGGGCGGCTTCTGTATCTGCAAGACCTGGTGG 1740
DB 1787 ATCCCGTGGAGATTTCCGGTACATCTGGGGCGGTTTGGCTATCTGACGACATGGTTG 1846
QY 1741 AGCTGACGCGCTCGCTGCTCAGCGGCGCAACCCCGGGCGCGCTCTACCTGACG 1800
DB 1847 AACAGGGATCACAAGAGGACGAGTGCAGGGCGGAGGCTCCAGTTGGAATCTACCTCCAGC 1906
QY 1801 AGATGCCCTATCCGTTGCTATGTGAGGAGGTTTCTCTGGTGTGCTGAGCGGTGCTGC 1860
DB 1907 AGATGCCCTATCCGTTGCTGAGGAGTTCCTTATGATCATCATCTGTAACCGCTGTTTCC 1966
QY 1861 CGCTCTTCTGACGCTGGCTGATCTACTCCGTGACACTGACAGTGAAGCGCGTGTGC 1920
DB 1967 CTATCTTCTGCTGAGTGGATCTACTCTGTCTCCATGACTGTGAAGACATCGTCT 2026
QY 1921 GGGAGAGGAGACGCGGTGGGGACACATGCGCGCCATGGGCTCAGCGCGCGGTGC 1980
DB 2027 TGGAGAGGAGTTGGCATGAAGGAGACCTTGAAAAATCAGGGTGTCTCCAAATGACGTGA 2086
QY 1981 TCTGGCTAGGCTGGTTCTCAGCTGCTGGGCGCTTCTGCTGAGCGCGCGGTGCTGG 2040
DB 2087 TTTGGTGTACCTGGTCTTGACAGCTTCTCATCATGTGATGAGCATCTTCTCTCTGA 2146
QY 2041 TTCTGGTCTCAAGCTGGGGACATCTCCCTACAGCACCCCGGGCGTGTCTTCTCTGT 2100
DB 2147 CGATATTCTATGATGAGAGATCTTACATTACAGGACCCATTCACTCTTCTCTGT 2206
QY 2101 TCTTGGAGCTTCCGCTGGGCGCACCGGTGACCCAGAGCTTCTGCTCAGCGCTTCTTCT 2160
DB 2207 TCTTGTGGCTTCTCCACTGCGACCATCATGTGTGTCTTCTGCTCAGCACCTTCTTCT 2266
QY 2161 CCGGGCCAACTGGCTGGGCTCGGGCGGCTGGCTTCTTCTGCTGAGCGCGGCTGCTGCCT 2220
DB 2267 CCAAGGCCAGTCTGGCAGCAGCTGTAGTGTGTCTATTTTCAACCTCTACCTGCCAC 2326
QY 2221 ACCTGCTGTGTGTGGGGGACCGGCTGCGCGGGTGGCGCGTGGCGCGAGCC 2280
DB 2327 ACATCTGTGTGTCCCTTGGCAGGACCCGATGACCGCTGAGCTGAAGAGCTGTGAGCT 2386
QY 2281 TGCTGTCCCGCTGGCTTCCGCTTGGCTGCGAGAGCTTGGCTTCTGCTGGAGGACGAG 2340
DB 2387 TACTGTCTCCGGTGGCATTTGGATTGGCACTGAGTACCTGTTGCTTGAAGAGCAAG 2446
QY 2341 GCGAGGGCGCGAGTGGACAACTGGGACACCGGCTTACGGCA---GACGTCTTACGCC 2397
DB 2447 GCCTGGGCTCAGTGGAGAACATCGGGAAACAGTCCACAGGAGGGGACGAAATTCAGCT 2506
QY 2398 TGGCCAGGCTCTCTGGCTTCTGCTGTGGAAGCGGCGCTCTAGGGCTCGCCACCTGGT 2457
DB 2507 TCTGCTCTCAGCAGATGATGCTCTTGTGATGCTGTCTATGGCTTACTCGCTTGT 2566
QY 2458 ACCTGGAAGCTGTGTGCCCCAGGCCAGTACGGGATCCCTGAAACCATGGAATTTTCTTTTC 2517

DB 2557 ACCTTGATCAGGTGTTTCCAGGAGACTATGGAACCCCACTTCTCTGGTACTTCTTCTTAC 2626
QY 2518 GGAGGAGCTACTTGGTGGC----- 2535
DB 2627 AAGAGTGTATTTGGCTTGGCGTGAAGGGTGTTCACACAGAGAAAGAGGCGCTTGAAAA 2686
QY 2536 -GACCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCACCCGCTGGACCCAAAGGTG--- 2591
DB 2687 AGACCGAGCCCTTACAGAGGAAACGAGGATCCAGAGACCCAGAGGAATACAGACT 2746
QY 2592 --CTGGTAGAAGAGCAACGCGCGGCTGAGTCTTGGGCTATCTTTCGACGCTTGAGA 2649
DB 2747 CTTCTTTGAACGCTGAGCATCCAGGGTGGTCTCTGGGGTATGCGTGAAGAACTCTGGTAA 2806
QY 2650 AGCGCTTCTTGGAAAGCCCGACGAGCCCTGCGGGGCTCAGCTGAGCTTCTACCAAG 2709
DB 2807 AGATTTTGAAGCCCTGTGGCGGCGAGCTGTGACCGTCTGAACATCACTTCTTACGAGA 2866
QY 2710 GCCATCATCACCGCTTCTTGGGCCCAACACGGGGCGGCAAGACCAACCACTTGTCCATCT 2769
DB 2867 ACCAGATCACCGCATTCTTGGGCCACATGAGCTGGGAAACCAACCACTTGTCCATCC 2926
QY 2770 TGAAGTGGCTCTTCCACACCGAGTGGTCTCTGCGCTTCACTCTGGGCGACGAGCTCGCT 2829
DB 2927 TGAAGGGTCTGTGCGACCAACCTCTGGGACTGTGCTCGTTGGGGAAAGGACATTGAAA 2986
QY 2830 CCAGCATGGCGCCCATCCGGCCCCACCTGGGGCTGTCTCTCAGTACAAACGCTGTTTG 2889
DB 2987 CCAGCTGGATGCACTCCGAGCAGCTTGTCCAGCATGTGTCCACAGCACAACTCTGTTCC 3046
QY 2890 ACATGCTGACCTGGACGAGCAGCTCTGGTTCTATGGGCGGTGAAGGTCTGAGTGCAG 2949
DB 3047 ACCACCTCACGGTGGCTGAGCACAATGTGTCTATGCCCAGCTGAAAGAAAGTCCGAG 3106
QY 2950 CTGTAGTGGGCCCCGAGCAGGACCGCTGTCTGTCAGGATGTGGGGTGGTCTTCAAAGCAGA 3009
DB 3107 AGAGGCCCACTGGAGATGGAAGCCATGTTGGAGGACACAGGCTCCACCAAGAGCGA 3166
QY 3010 GTGTGCACTCGCCACCTCTCTGGTGGATGCAACGGAAGCTGTCTCGTGGCATTGCT 3069
DB 3167 ATGAAGGCTCAGGACCTATCAGGTGGCATGCAAGAAAGCTGTCTGCTTGGCATTGCT 3226
QY 3070 TTGTGGGCGGCTCCCAAGTGTATCTTGGACGAGCTTACGCTGGCTGGTGGATCTGCTT 3129
DB 3227 TTGTGGGAGTCCAAAGTGTGTATCTGGAGAACCCACCTCTGGGGTGGACCTTACT 3286
QY 3130 CCGCGCGGCTTATTTGGAGCTGTCTGCTCAAAATACCGAAGAGTGCACGCTGATCTCT 3189
DB 3287 CGAGACGCTCAATCTGGGATCTGCTCTGAAAGTATCGCTCAGGCAAGAACCATCATGT 3346
QY 3190 CCACCCACCACTGGATGAGCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTG 3249
DB 3347 CCATCCACCATGGACGAGCCGACCTCTCTGGGAGCCGCAATTCGCAATTCATTCGCCAGG 3406
QY 3250 GCGCTTGTGTGTGTGGTCTCCCACTCTTCTCTGGCGCGTCACTGCTGGGCTCGGCTACT 3309
DB 3407 GAAGGCTCTACTGTCTAGGCAACCCCACTCTTCTTGAAGAACTGTCTTGGCAGAGCTTGT 3466
QY 3310 ACCTGACGCTGGTG-----AAGGCCCGCTTGGCCCTGACCAACCAATGAGAAAG 3357
DB 3467 ACTTAACTTGGTGGCAAGATGAATAACATCCAGAGCCAAAGAAAGGAGTGAAGGGA 3526
QY 3358 CTGACACTGACATGGAGGGCAGTGTGGACACCAAGC-----AGAAAGAAAGA 3405
DB 3527 CCTGAGCTGTCTGTAAAGGGTTTCTCCACAGCTGTCCAGCCCAAGCTGATGACCTAA 3586
QY 3406 ATGGCAGCCAGGGCAGCAGAGTCCGCACTCTCTCAGCTGTGCGGCTGTGACAGACTGG 3465
DB 3587 CTCAGAAACAAAGTCTGTGATGGGATGTAAATGAGCTGATGATGTAGTTCTCCACCAATG 3646
QY 3466 TGCCCGGGGCAAGCTGTGGAGGAGCTGCCACAGAGCTGGTGTGCTGCTGCCCTTACA 3525

Db 3647 TTCAGAGGCAAGCTGGTGGAGTGCAATTGGTCAAGAACTTATCTCTCTCTTCCAAATA 3706
Qy 3526 CGGGTCCCATGATGCGGAGCTTGGCCACACTCTTCCGAGAGCTAGACACCGGCTGGCG 3585
Db 3707 AGAATCTCAAGACACAGACATATGCCAGCCCTTTTCAGAGAGCTGGAGAGACGCTGGCTG 3766
Qy 3586 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCTCGAGGAGATCTTCTGAAGG 3645
Db 3767 ACTTGGTCTCAGCAGTTTGGAAATTTCTGACACTCCCTCGGAAGATTTTTCGAGG 3826
Qy 3646 TGGTGGAGAGTGTGTGCGGACACAGATATGGAGGATGCCAGCTGGCG- - - - -CAGC 3699
Db 3827 TCAGGAGGATCTGATTACAGGACCTCTGTTGCGGGTGGCGCTCAGCAGAAAAAGAAA 3886
Qy 3700 ACCTATGCAAGGATGTCTGGCTAGACGTAAACCTTGGCGCTCAAGATCCGCGCACAGG 3759
Db 3887 ACGTCAACCCCGACACCCCTGCTGGTCCGAGAGAAAGGCTGGAAGACACACCCAGG 3946
Qy 3760 AGACAGGCTGGAGAACGGGGAACCAAGCTGGGTGAGCCCGAGAGACTGACACAGGCTCTG 3819
Db 3947 ACTCCAATGTCTGCTCCGAGGGGCGGCTGCTCACCCAGAGGGCCAGCTCCCCAG 4006
Qy 3820 GGCCAGACGCGTGGGCGCGGTAC- - - - -AGGCTGGGCACTGACCCCGCAGAGCTCC 3873
Db 4007 AGCCAGAGTCCCGAGGCGCCGAGCTCAACACGGGACACAGCTGGTCTCCAGCATGTC 4066
Qy 3874 AGGCCGTGCTTCAAGCGCTTCTGCTTGGCGCGCGCAGCGCGCGGCGCTTGGCCC 3933
Db 4067 AGGCGCTGCTGCTCAAGAGATTTCCACACCACTCCGCGACCCCAAGGACTTCTCGCGC 4126
Qy 3934 AGATCGTGTGCTGCGCTCTTTGTGGCGTGGCGCTCGGTTCAGGCTCATCGTGCTC 3993
Db 4127 AGATCGTGTGCTGCGCTTACCTTTGTGTTTGTGCTCTGATGCTTTCTATTGTTATCCCTC 4186
Qy 3994 CTTTGGGCACTACCCGCTCTGCGGCTCAGTCCCACTATGACGCTGCTCAGGTGCT 4053
Db 4187 CTTTGGCGAATACCCGCTTTGACCTTCAACCCCTGGATATATGGGACAGTACACCT 4246
Qy 4054 TCTTCAGTAGAGACGCGCCGAGGAGCCCTGGAGCGTGCGCGGCTGCTCGAGGCGTGCTGC 4113
Db 4247 TCTTCAGCATGGATGAACACGAGGAGTGAGCAGTTTCACGTACTTGCAGACGCTCCTCTGA 4306
Qy 4114 AGAGGACGAGCTGGAGAGCGCCCGCTGACGCA- - - - - 4147
Db 4307 ATAAGCCAGGCTTTGGCAACCGTGTCTGAAGAGGGTGGCTTCGAGTACCCCTGTG 4366
Qy 4148 -TAGCTCCACAGGTTCTCGGCACACAGAGTTCTGTGTAAGTGGCCAAAGTCTTGGCCA 4206
Db 4367 GCAACTCAACACCTGGAAGACTCTTCTGTGTCCCAACATCACCCAGCTGTTCAGAG 4426
Qy 4207 GTGGCAACTGGAACCCAGAGTCTCCATCCCGAGCTGCCAGTGTAGCCAGCCCGGTGCC 4266
Db 4427 AGCAGAAATGGACACAGGCTCAACCTTTCAGCATCTGCAAGTGTGAGCAGCAGGAGAAC 4486
Qy 4267 GCGCCTGTGCGGACTCGCGCTGACGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4326
Db 4487 TCACCATGCTGCCAGAGTGCCCGAGGGTGGCGGGGCTTCCCGCCCGCCAGAGAAC 4546
Qy 4327 CCGGCTCTGGGAGTGGTTTCAAGACCTGACAGCGCCGGAACCTGTCTGACTTCTCGTCA 4386
Db 4547 AGCGCAGACGGAATTTTACAGACCTGACGGACAGGACATCTCCGACTTCTTGGTAA 4606
Qy 4387 AGACTTACCGCGCTGTGTGGCCAGGGCTGAAAGCTAAGAGTGGGTGAATGAGTCA 4446
Db 4607 AAACGTATCTGCTCTTATAAGAGAGCAGTCTAAAGAGCAAAATTTCTGGGTCAATGAACAG 4666
Qy 4447 GGTACGAGGCTTCTCGCTGGGGGCGGAGACCCAGGCTGCGC- - -CTCGGGCCAGAGT 4503
Db 4667 GGTATGAGGAATTTTCATTTGAGAGAAAGCTCCCAAGTCTGCTGCTGCTGCTGCTGCTG 4726
Qy 4504 TGGCGCTCAGTGGAGAGTGTGGCGCTGTGAGTCCCTGCTGCTGCTGCTGCTGCTGCTG 4563
Db 4727 TTGTTGGGTTTTTAAGCGACTTGGCCGGATCATGATGTGAGCGGGGCGCTTACTACTA 4786

Qy 4564 ACCGTGCTCTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCA 4623
Db 4787 GAGAGCCCTCTAAGAAATACCTGATTTCTTTAAACATCTAGAAATCTGAAGACACATTA 4846
Qy 4624 AGATCTGGTTCAACAAACAAAGGCTGGCACTTCCATGTGTGGCTTTGTCAACCGAGCCAGCA 4683
Db 4847 AGGTGTGGTTTAAATAACAAAGGCTGGCATGCGCTGTGCTCAATGTGTGCCACA 4906
Qy 4684 AGCAATCTCTCGTGTCTACCTGCGCCCGAGGCGCGCCGCGCCAGCCCAAGCATCAACA 4743
Db 4907 ACGCCATCTTACGGGCGAGCTGCTTAAGACAGAGGAGCCCGAGGAGTATGGAATCACCG 4966
Qy 4744 CACTCAACACCCCTTGAACCTCACCAAGAGCAGCTGTCTGAGGCTGCACTCATGSCCT 4803
Db 4967 TCATTAAGCAACCCCTGAACCTGACCAAGAGAGCTCTCAGAGATTAACAGTCTGACCA 5026
Qy 4804 CTTGGTGGAGCTCTCGTCTCCATCTGTGTGTCTTTGCGCATGTCTTTGTCCCGGCA 4863
Db 5027 CTTTCAGTGGATGCTGTGGTTGCCATCTGCGTGATTTTCTCCATGTCTTCTGCTCCAGCA 5086
Qy 4864 GCTTCACTCTTGTCTCATTTGAGGAGCGAGTCAACCGAGCAAGCACTGCGAGCTATGG 4923
Db 5087 GCTTTGCTCTTATTTTGTATCCAGAGCGGGTGAACAAATCCAAAGCACCTCCAGTTATCA 5146
Qy 4924 GGGGCTGTCCCGCACCTCTACTGCTTGGCAACTTTCTCTGGGACATGTGTAACTACT 4983
Db 5147 GTGGAGTGAAGCCCAACACTACTGCTGTGACCACTTCTCTGGGACATCATGAATATT 5206
Qy 4984 TGGTGCCAGCATCATCGTGGTCTCATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGG 5043
Db 5207 CCGTGAAGTCTGGGCTGGTGGGCACTCTTCATCGGTTTTCAGAAAGACCTACACTT 5266
Qy 5044 CCCTGCCAACCTGCT 5103
Db 5267 CTCAGAAACCTTCT 5326
Qy 5104 CGCTCATGTACCCAGCCT 5163
Db 5327 CCATGATGTACCCAGCATCT 5386
Qy 5164 CTGCAATAAACCTCTTTTATTTGGCATCAATGGAGCATGCGCACCTTTTGTCTTTGAGCTCT 5223
Db 5387 CTTGTCTAATCTGTTCATCGGCATCAACAGCAGTGTCTATTACCTTCATCTTGAATTA 5446
Qy 5224 T- - -CTCTGATCAGAACTGCGAGGAGTGAACCGGATCTTTGAACAGGCTCTCTTATCT 5280
Db 5447 TTGAGAAATAACCGGACGCTGCTCAGGTTCAACCGCGTGTGTGAGAAAGCTGCTCATTTGCT 5506
Qy 5281 TCCCCCACTTCTGTTGGCGCGGGCTCATTTGACATGTTGCGGAACCGAGCCCATGGCTG 5340
Db 5507 TCCCCCACTTCTGCTGGCGCGGGCTCATTTGACCTGACCTGACCTGAGCAGGCTGTGACAG 5566
Qy 5341 ATGCTTTGAGCCTTTGGGAGACAGGAGTTCAGTCAACCCCTGCGCTGGGAGGTGTGCG 5400
Db 5567 ATGCTATGCCCGTTTGGTGGAGGACCTCTGCAATCGTTTCCACTGGGACCTGATTG 5626
Qy 5401 GCAAGAACTCTTTGGCCATGTTGATACAGGGGCGCTCTTCTCTCTCTCTCTCTCTCTCTCT 5460
Db 5627 GGAAGAACTGTTTGGCATGTTGGTGAAGGGGTGTGTACTTCTCTCTCTCTCTCTCTCTCTCT 5686
Qy 5461 TGCGACACCGAAGCAACTCTCTCCACAGCCAGGCTGAGTCTCTGCGCACTCTCTGGAG 5520
Db 5687 TCAGGCGCACTTCT 5746
Qy 5521 AGGAGGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGGTTCAGAGGAGCCACCCAGGGG 5580
Db 5747 ATGAAGATGATGTGGCTGAAGAAAGACAAAGAAATATTACTGGTGGAAATAAAACATG 5806
Qy 5581 ATGTGTTGGTGTGAGGAACTTGAACAGGTATACCGTGGGAGAGAGATGCCAGCTGTGTTG 5640
Db 5807 ACATCTTAAGGCTTACATGAATAACCAAGATTTTATCTGGGCACTCTCCAGCCCGAGCTGG 5866

Qy	5641	ACCGTGTGTGCTGGGGAATCCCCCTCGGTGAGTGTFTTTGGGCTGCTGGGTGTGAATGGAG	5700
Db	5867	ACAGGCTGTGTGTTCGGAGTTTCGCCCTCGAGAGTGTCTTTGGCCTCTCTGGGAGTGAATGGTG	5926
Qy	5701	CAGGGAAAGCGTCCACGCTTTTCGCATGTGTGCGGGGACACATTTGGCCAGCAGGGGCCAGG	5760
Db	5927	CCGGCAAAACAAACACATTCAGATGCTCTACTGGGGACACACAGTGCACCTTCAGGGGATG	5986
Qy	5761	CTGTGCTGGCAGGCCACAGCGTGGCCGGGAACCCAGTGTGTGGGCACCTCAGCATGGGAT	5820
Db	5987	CCACCGTAGCAGGCAAGAGTATTTTACCAATATTTTCTGAAGTCCATCAAAATATGGGCT	6046
Qy	5821	ACTGCCCTCAATTCGATGCCATCTTTGAGCTGTGTGACGGGCCGCGAGCACTCGAGCTGC	5880
Db	6047	ACTGTCTCAGTTTGTATGTCAATCCATGAGCTGCTCACAGCAGCAGAAACATCTTTACCTTT	6106
Qy	5881	TTGGCGCCCTGCGCGGTGTCCCGGAGGCCAGGTTGCCCCACCGCTGGCTCAGGCCCTGG	5940
Db	6107	ATGCCCGGCTTCGAGGTGTATCCAGCAGAAAGAAATTCGAAAAGGTTTGCNAATCGGAGTATTA	6166
Qy	5941	CGCGTCTGGGACTCTCATGTTACGCAGACCGGCTTCGAGGCACCTACAGCGGAGGGAAACA	6000
Db	6167	AGAGCTTGGGCTGACTGTCTACGCCGACTGCTCGGTGGCACTGACAGTGGGGGCAACA	6226
Qy	6001	AAGCAAGCTGGCAGCGCCCTTGCGCTGTGTGGGACCCAGCCGTGGTGTCTTCGTGACG	6060
Db	6227	AGCGGAACTCTCCACAGCCATGSCACTCATTTGGCTGCCCAACCGCTGCTGCTGGATG	6286
Qy	6061	AGCGACACAGGATGGAACCCAGCGCGGGCGCTTCCTTTGGAAACAGCCTTTTGGCCG	6120
Db	6287	AGCCACCAACAGGATGGACCCCGAGCAGCGCGCATGCTGTGGAAACGTATCGTGGACGA	6346
Qy	6121	TGTTGGGAGGGCCGTTTCAGTGATGCTCACTCCCATAGCATGAGGAGGTGTGAAGCCG	6180
Db	6347	TCATCAGAGAGGGAGGGCTGTGGTCTCACAATCCCAAGCATGGAAGAATGTGAGCAC	6406
Qy	6181	TCGTCTCGCCCTAGCCATCATGGTGAATGGGGGGTTCCGCTGCTGGGCAGGCCCGCAAC	6240
Db	6407	TGTTGTAACCGGCTGGCCATCATGGTAAGGGCGCCTTTCAATGTATGGGCACCATTCAGC	6466
Qy	6241	ATCTCAAGGCGAGATTCCGGCGGGTCAACA CTGACCTCTCGGGTGCCTCCCGCCGCAAGGT	6300
Db	6467	ATCTCAAGTCCAAATTGGAGATGGCTATATCGTCAATGAAGATCAAAATCCCCGAAG	6526
Qy	6301	CCAGCGGCGAGCGGCCCTT-----CGTGGCGCCGAGTTCCCTGGGTGGAGCTGGCG	6354
Db	6527	ACGACTGTCTCTGACCTGAACCTCTGTGGAGCAGTTCTTTCAGGGGAACTTCCCAAGCA	6586
Qy	6355	AGGCATGAGGCGCGCTCGCTTCAGCTGCCG---CGGGAGGGCGCTGCGCCCTCG	6411
Db	6587	GTGTGCAGAGGAGAGGCATCAACAATGCTTCAGTTCAGGTCTCGTCTCTCTCCCTCG	6646
Qy	6412	CGCGCGCTTTTGGAGAGCTGGCGGTGCACGCGCGCAGAGCACGGCGTGGAGGACTTTTCCG	6471
Db	6647	CGAGGATCTTCCAGTCTCTCTCTCCACAAAGACAGCCCTGCTCATCGAGGAGTACTCAG	6706
Qy	6472	TGAGCCAGACGATGTGGAGGAGGTATTTTGTACTTCTCCAAAGGACCAG	6521
Db	6707	TCACACAGACCACTCGACAGGTGTTTGTAAATTTTGTCTAAACAGCAG	6756

RESULT 3

US-09-032-438C-1
; Sequence 1, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael

Db 4415 TACCGAGGATTCGATTCAGGACCTCTGTTTTCGGGTGGCGCTCAGCAGAAAAGAGAAA 4474
Qy ACCTATGACACAGGCAATGCTGGCCCTAGACGTAAACCTCTGGCTCAAGATGCGCCGACACAGG 3759
Db 4475 ACGTCAACCCCGACACCCCTGCTTGGGTCCAGAGAGAGGCTGGACAGACACCCAGG 4534
Qy 3760 AGACAGCGCTGGAGAAACCGGGGAAACAGCTGGGTGAGCCCGAGAGACTGACCAAGGCTCTG 3819
Db 4535 ACTCAATGTCTGCTCCCGAGGGGCGGGCTGCTCAACCAAGAGGGCGAGCTCCCGCAG 4594
Qy 3820 GSCCAGACGCGTGGGCGGGGTAC-----AGGGCTGGGCACTGACCCGCGCAGAGCTCC 3873
Db 4595 AGCCAGAGTGCCAGGCGCCGAGCTCAACACGGGACACAGCTGTGCTCTCAGCATGTGC 4654
Qy 3874 AGGCCCTGCTTCTCAAGGCTTCTTGCTTGGCCGCGCAGCGCGCGGCTGTTGCGCC 3933
Db 4655 AGGCGCTGCTGGTCAAGAGATTCACACACACCATCCGACGACCAAGAGACTTCTGGGCG 4714
Qy 3934 AGATCGTGTGCTGCTCCCTCTTTGTGGGCTGGCCCTCGTGTGAGCCTCAFCGTGCTC 3993
Db 4715 AGATCGTGTGCTGCTCCGCTACCTTGTGTTTGGCTCTGATGCTTCTATTTATCCCTC 4774
Qy 3994 CTTTTCGGGCACTTACCGGCTCTGCGGCTCAGTCCACCATGTACGGTGTCTCAGGTGCT 4053
Db 4775 CTTTTCGGGCAATACCCCGCTTTCACCTTCACCCCTGGATATATGGGCGAGTACACCT 4834
Qy 4054 TCTTCAGTGAAGACGCCCCAGGGGACCTGGACGTTGCCGCTGTGCGAGGCGTGTGTC 4113
Db 4835 TCTTCAGCATGGATGAACAGGCGAGTGAAGTTCACGGTACTTGCAGACGCTCTCTG 4894
Qy 4114 AGGAGGAGGACTGGAGGAGCCCCAGTGCAGCA----- 4147
Db 4895 ATAAGCAGGCTTGGCAACCGCTGCTGTAAGGAGGGTGGCTTCCGAGTACCCCTGTG 4954
Qy 4148 -TAGTCCCAAGGTTCTCGGACACCAAGTTCCTGCTGAAGTGGCCCAAGTCTTGGCCA 4206
Db 4955 GCAACTCAACACCTTGGGAAGACTCCTCTGTGTCCCCAAACATCACCCAGCTGTTCCAGA 5014
Qy 4207 GTGGCACTGACCCAGAGTCTCCATCCAGCCTGCGAGTGTAGCCAGCCCGTGGCC 4266
Db 5015 AGCAGNAATGACACAGTCAACCTTCAACATCTGCGAGTGCAGCACAGGAGGAGC 5074
Qy 4267 GCGCCTGCTGCCGACTGCCCGGCTGAGCTGTGGTCCCCCTCCGCCCCAGGAGTGA 4326
Db 5075 TCACCATGCTGCAGAGTGCCCGAGGGTGGCGGGGCTCCGCGCCCCCAGAGAACAC 5134
Qy 4327 CCGGCTCTGGGAGTGTTCAGAACCTGACAGGCGGAACTGTCTGACTTCTGTGTC 4386
Db 5135 AGCGCAGCACGGAATTTCAAGAACCTTGACGGGACAGGAACATCTCCGACTTCTTGGTAA 5194
Qy 4387 AGACTTACCCGCGCTGTGCGCCAGGCGCTGAAGACTGAAGTGGGTGAATGAGGTCA 4446
Db 5195 AAAAGTATCTGCTTATAGAGAGCAGCTTAAGAGCAAAATTCGGGTCAATGAACAGA 5254
Qy 4447 GGTACGGAGGCTTCTGCTGGGGGCGCGAGACCCAGGCTGCCC---CTCGGGCCAAAGAGT 4503
Db 5255 GGTATGGAGGAATTTCCATTGGAGAAAGCTCCAGTCGTCCCCATCACCGGGGAGGCAC 5314
Qy 4504 TGGGCGCTCAGTGAGAGAGTTTGGGCGCTGTGAGTCCCTGCTGCGGGGGCGCTCG 4563
Db 5315 TTGTTGGGTTTTTAAGCGACCTTGGCCGGATCATGAATGTGAGCGGGGCGCTTACATA 5374
Qy 4564 ACCGTGCTCTGAAAACTCAGCCTGGGCTCAGCGCTGGATGCTCAGGACAGTCTCA 4623
Db 5375 GAGAGGCTCTTAAGAAATACCTGATTTCTTTAAACATCTAGAAACTGAAGAACATTA 5434
Qy 4624 AGATCTGTTCAACAAAGGCTGGCACTCCATGGTGGCTTTGTCAACCGAGCCAGCA 4683
Db 5435 AGGTGTGTTTAAATAACAAAGGCTGGCATGCCCTGGTCACTTCTCAATGTGGCCNACA 5494
Qy 4684 AGCMACTCTCGTGTCTCACTGCCCCCAGGCGCGGCCCCGACCGCCACAGCATACCA 4743

Db 5495 AGGCATCTTACGGGCGACGCTGCTTAAGGACAGGAGCCCCGAGGATATGGAATACCG 5554
Qy 4744 CACTCAACACCCCTTGAACCTCAACAGGAGCAGCTGCTGAGGCTGCACCTGATGGCCT 4803
Db 5555 TCATTAGCAACCCCTGAACCTGACCAAGGAGCAGCTCTCAGAGATTACAGTGTGACCA 5614
Qy 4804 CTTGGTGGACGCTCTGCTCTCATCTGTGTGGTCTTTGGCCATGTCCTTTGTCCCGGCA 4863
Db 5615 CTTCAGTGGATGCTGTGGTGGCCATCTGCGTGAATTTCTCCATGTCCTTCTGTCGCCAGCA 5674
Qy 4864 GCTTCACTCTTGTCTCATTTGAGGAGGAGTCAACCGAGGCAAGCACCTGCGAGCTCATGG 4923
Db 5675 GCTTTGCTCTTTATTTGATCCAGGAGGGGTGAACAAATCCAAGCACCTCCAGTTTATCA 5734
Qy 4924 GGGGCTGTCCCCCAACCTCTACTGGCTTGGCAACTTCTCTGGGACATGTGTAACACT 4983
Db 5735 GTGGAGTGAGCCCAACCACTACTGGGTGACCACTTCTCTGGGACATCATGAATTAT 5794
Qy 4984 TGGTGCCAGCATGATCGTGTGCTCATCTTTCTTGGCCCTCCAGCAGAGGCAATATGG 5043
Db 5795 CCGTAGTGTCTGGGCTGTGGTGGGCATCTTTCATCGGGTTTCAGAAAGAACCTACACT 5854
Qy 5044 CCGTCCCAACCTGCTGCTCTGCTGTGTGCTACTACTGTATGGCTGGTGCATCACAC 5103
Db 5855 CTCCAGAAAACCTTCTGCTGCTTGTGGCAGTCTCTGCTGTATGGATGGGCGGTCAATTC 5914
Qy 5104 CGCTCATGTACCCAGCCTCTTCTTCTCCGTCGCCAGCACAGCCTATGTGGTGTCTCA 5163
Db 5915 CCATGATGATCCAGCATCTTCTGTTGATGTCCCGCAGCACAGCCTATGTGGCTTAT 5974
Qy 5164 CCTGCATAAACCTCTTTATGGCATCAATGGAAGCATGGCCACCTTTGTGCTTGAGCTCT 5223
Db 5975 CTTGTGCTAATCTGTTCATCGGCATCAACAGCAGTGTATTACCTTCATCTTGAATAT 6034
Qy 5224 T---CTGATCAGAACTGCGAGGAGTGAGCGGATCTTGAACAGAGTCTTCTTATCT 5280
Db 6035 TTGAGAAATAACCGGACGCTGCTCAGGTTCAACCGCGTGTGAGGAAGCTGCTCATTTGCT 6094
Qy 5281 TCCCCCACTTCTGCTTGGGCGGGGCTCATTTGACATGTTGCGGAACCCAGGCAATGGGTG 5340
Db 6095 TCCCCCACTTCTGCTTGGGCGGGGCTCATTTGACCTTGTGCACTGAGCCAGGCTGTGACAG 6154
Qy 5341 ATGCTTTGAGCCTTTGGGAGACAGGAGTTCAGTCAACCTCGCTGGTGGAGGTGGTGG 5400
Db 6155 ATGCTATGCGCGTGTGGTGAGGAGCCTCTGCAAAATCCGTTCCACTGGGACCTGATTG 6214
Qy 5401 GCAGAACCTCTTGGCCATGCTGATACAGGGGCCCCCTTCTCTCTTCACTACTGCTGC 5460
Db 6215 GGAAGAACCTGTTTGGCCATGGTGGTGAAGGGGTGGTGTACTTCTCTGACCCCTGCTGG 6274
Qy 5461 TGCAAGCACCGAAGCAACTCTCTGCGCACAGCCCGGTGAGGTCTCTGCACTCTCTGGAG 5520
Db 6275 TCCAGCGGCATCTTCTCTCCCAATGGATTGCGGAGCCCACTAAGAGGCCCATTTGTTG 6334
Qy 5521 AGGAGGACGAGGATGTAGCCCGTGAACGGGAGCGGGTGTCCAAAGGAGCCACCCAGGGGG 5580
Db 6335 ATGAAGATGATGATGTGGCTGAAGAAAGACAAGAAATTTACTTGGTGAATAAACAATG 6394
Qy 5581 ATGTTGGTGTCTGAGGAACCTTGACCAAGGTATACCGTGGGAGAGGATGCCAGTGTG 5640
Db 6395 ACATCTTAAAGCTACATGAACATAACCAAGATTTATCCAGGCACTCCAGCCAGCAGTGG 6454
Qy 5641 ACCGCTTGTGCTGGGATTTCCCGCTGTGAGTGTGTTGGGCTCTGCGGTGTGAATGGAG 5700
Db 6455 ACAGGCTGTGTGCGAGTTGCGCTGTGAGAGTGTCTTTGGCTCTTGGGAGTGAATGGTG 6514
Qy 5701 CAGGAAAGAGCTCCACGTTTTCGATGTTGAGCGGGGAGACATTTGGCCAGCAGGGGCGAGG 5760
Db 6515 CCGCAAAACACACCATTCAGATGTCTACTGGGAGACACCAAGTACCTCAGCGGGATG 6574
Qy 5761 CTGTGCTGGAGGCGCACAGCGTGGCCCGGGAACCCAGTGTGCGGCACTCTAGCATGGAT 5820
Db 6575 CCACCGTAGCAGGCAAGAGTATTTTAAACCAATATTTCTGAAGTCCCATCAAAATATGGGCT 6634

QY 5821 ACTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGCCGCGAGCACTGGAGCTGC 5880
DB 6635 ACTGCTCTCAGTTGATGCAATCATGATGCTGCTCAAGAGCAGAAATCTTTTACCTTT 6694
QY 5881 TTGCGCCCTCGCGGCTGCTCCGAGGCCCAGGTTGCCAGACCGCTGGCTCAGGCTCG 5940
DB 6695 ATGCCCGCTTCGAGGTGTACACAGCAAGAAATCGAAAGTTGCAAACTGGAGTATTA 6754
QY 5941 CGGCTCTGGAGCTCTCATGTTACGACAGCCGGCTGTCAGGCACTTACAGCGAGGGAACA 6000
DB 6755 AGAGCTCTGGGCTGACTGTCTACGCCGACTGCTGGCTGCGACGCTACAGTGGGGCAACA 6814
QY 6001 AACGCAAGCTGGCAGCGGCTGCGCTGCTGGGACCCAGCCGCTGGTGTCTGACG 6060
DB 6815 AGCGAAACTCTCCACAGCCATGCACTCAATTTGGCTGCCACCGCTGGTGTCTGTAAG 6874
QY 6061 AGCGACACACAGGATGGAACCCAGCGCGGGCTTCTTTTGAACAGGCTTTTGGCCG 6120
DB 6875 AGCCACCAAGGATGGAACCCAGGCAAGCGCATGCTGTGGAAGCTCATCGTGAGCA 6934
QY 6121 TGGTGGGAGGCGCTTCTAGTGTATGCTACCTCCCATAGATGAGAGAGTGTGAAGCG 6180
DB 6935 TCATCAGAGAAGGAGGCTGTGCTCTCATATCCACAGCATGGAAGATGTGAGCAC 6994
QY 6181 TCTGCTCGGCTAGCCATCATGTTGAATGGCGGTTCCGCTGCTGGCAGCCGCAAC 6240
DB 6995 TGTGTACCCGCTGGCCATCATGTTAAAGGGCGCTTTTCGATGATGGCACCATTTCAG 7054
QY 6241 ATCTCAAGGGCAGATTCGCGCGGGTTCACACATGACCTGCGGGTCCCGCGCAAGGT 6300
DB 7055 ATCTCAAGTCCAAATTTGGAGATGGCTATATGCTCAAAATGAAGATCAATCCCGNAGG 7114
QY 6301 CCAGCGCGAGCGGCTT-----CGTGGCGCGAGTTCCCTGGGTGCGAGTGGCGG 6354
DB 7115 ACGACTGCTTCTGACCTGAAACCTGTGGAGCAGTTCTTCCAGGGGAACCTTCCAGGCA 7174
QY 6355 AGCAGATGAGGCGGCTGCTTCCAGCTGCG-----CGGAGGCGGCTGGCGCTGG 6411
DB 7175 GTGTGCAGAGGAGAGCACTAACATGTTCCAGTTCCAGGTCTCTCTCTCTCTCTCTG 7234
QY 6412 CGGCGCTTTTGGAGAGCTGGCGGTGACCGCGCAGAGCAGCGGCTGGAGGACTTTTCCG 6471
DB 7235 CGAGGATCTTCCAGCTCTCTCTCTCCCAAGGACAGCTGCTCATCGAGGATCTCAG 7294
QY 6472 TGAGCCAGAGATGCTGGAGAGATTTCTTTGACTTCTTCAAGGACAG 6521
DB 7295 TCACACAGACCACTGGACAGGTTTGTAAATTTTGTAAACAGCAG 7344

RESULT 4

US-09-032-438C-5
; Sequence 5, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C

; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 6705
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-032-438C-5

Query Match 19.0%; Score 1291.4; DB 4; Length 6705;
Best Local Similarity 53.9%; Pred. No. 9.5e-261;
Matches 3115; Conservative 0; Mismatches 2461; Indels 207; Gaps 14;

QY 859 TGAAGTTGCTGTGACAGGCGCTCTGAGAGTGTCAAGGACCTAGCAGCAGAGTGGGCCCCCT 918
DB 947 TGGGCATCTGTCTGACCTCTCTGTGGCTTACCCGAGGAGGTGGCTCTCGGGTGTCTCT 1006
QY 919 CCCTCAACTGGTACGAGGCTAGTACCTGATGGAGCTGGTGGGCGAGGCCAGAGATCCG 978
DB 1007 CCTCAACTGGTATGAAGACAATACTATAAGCCCTTCTTGGGGATTTGACTCCACAAGGA 1066
QY 979 CCCTGCC-----AGACAGCAGCCTGAGCCCGCTCTCTCGGAGCTGATTTGGAG 1026
DB 1067 AGGATCCTATCTATTCTTATGACAGAAGAACATCTTTTCTAATGATTGATCCAGA 1126
QY 1027 CCTGACAGCCACCGCTGTCCGCTGCTCTGGAGAGCCCTGGAAGCTCTGATCTCTG 1086
DB 1127 GCCTGAGTCAAACTCTTTAAACCAAAATCGCTTGGAGGGCGGCAAGCCCTTTGCTGATGG 1186
QY 1087 GGAAGTACTCTTTGACACAGATACACCTTTTACCGGAGCTCATGGCCCGAGTGAACC 1146
DB 1187 GAAAAATCTGTACATCTCTGATTCACCTGACAGACAGAGATCTTGAAGATGCCAACT 1246
QY 1147 GGAACCTTTCAGGAGCTCACCTCTCTGAGGATGTCCGGAGGTGTGGGAGATGCTGGGAC 1206
DB 1247 CAATTTTGAAGAACTGGAACAGCTTAGGAAGTTGTCAAAGCCTGGGAAGAGTAGGGC 1306
QY 1207 CCGGATCTTACCTTCATGAACGACAGTTCATATGTCCTATGTCGAGCGGCTCTGCTG 1266
DB 1307 CCAGATCTGCTACTTCTTTTGAACACAGCACACAGATGAACATGATCAGAGATACCCCTGG 1366
QY 1267 AGATCAGGATGAAGAGAGAGCAGCCAGACCTGAGGCGGGGACCAT-----GG 1320
DB 1367 GGAACCAACAGTAAAGAGCTTTTGAATAGGAGCTTGTGTGAAGAGGTATTTACTGCTG 1426
QY 1321 AGCCCTTGGCATCTTTTCTGGACCTTGGGAGCGGTGGCTACAGCTGGCAGGACGACACG 1380
DB 1427 AAGCCATCTTAACTTCTCTACAGGGCCCTCGGAAAGCCAGGCTGACGACATGGSCCA 1486
QY 1381 CTGATGTGGGCACTGTGGGGCAACGCTGGGCGGAGTGACGGAGTGGCTGTCTCTTGGACA 1440
DB 1487 ACTTCGACTGGAGGACATATTTAAACATCACTGATCGCACCTTCCGCTCGTCAATCAAT 1546
QY 1441 AGCTGGAGGCGGCAACCTCAGAGGAGCCCTGTGTGCGGGGCCCTGCAACTGTGCGCG 1500
DB 1547 ACCTGGAGTGTGGTCTCTGGATAAGTTTGAAGCTCAATGATGAACCTCAGCTTACCC 1606
QY 1501 AACATCGATTTCTGGCGGCGTCTCTTTCTTGGGACCTGAGGACTCTTTCAGAGCCCCACAG 1560
DB 1607 AAGTGGCCCTCTCTCTACTCGAGGAAACATGTTCTTGGGCGGAGTGGTATTCCCTGACA 1666
QY 1561 AGACCCCAACCCAGACCTGGGCGGCGGCAACGCTGGGCATCAAAATCCCGATGACATTTG 1620
DB 1667 TGTATCCCTGGACAGCTCTCTTACCAACCCCAACGCTGAAGTATAAGATCCGAATGGACATAG 1726
QY 1621 ACCTGTCTCAGGAGCAATTAAGATCAGGACAGGTTTGGGACCTTGGGCGGCGGCGCG 1680
DB 1727 ACCTGTGGAGAAACCAATGAATTAAGACAGGATTTGGGATTTCTGTTCCAGAGCTG 1786
QY 1681 ACCCTCTGACCGACCTGCGCTAGTGTGGGCGGCTTCTGTTGACTCTGCAAGACCTGGTGG 1740

Db 1787 ATCCCGTGAAGATTTCCTGGGTACATCTGGGGGGGTTTGCCTATCTGACGAGCATGGTTG 1846
Qy 1741 AGCGTGACCGCTCCGCGTGTCTCAGCGGCCAACCCCGGGCCGCGCTCTACCTGCGAGC 1800
Db 1847 AACAGGGATCAAAAGGAGCAGGTGCGAGCGGAGGCTCCAGTTGGAATCTACCTCCAGC 1906
Qy 1801 AGATGCCCTATCCGTTCTATGTGAGGACGTTGTTCTGCGTGTGCTGAGCGGGTCTGCGTC 1860
Db 1907 AGATGCCCTATCCGTTCTATGTGAGGACGTTGTTCTATGATCATCTCTGAACCGCTGTTTCC 1966
Qy 1861 CCCTCTTCTGACGCTGGCTCGATCTACTCTCGTGACACTGACAGTGAAGGCGGTGGTGC 1920
Db 1967 CTATCTTATGCTGTGAGTATGATCTACTCTCTCATGACTGTGAAGACATCGTCT 2026
Qy 1921 GGGAGAAGGAGACGCGCTGCGGACACCATCGCGCCATGAGGGCTCAGCGCGGGTGC 1980
Db 2027 TGGAGAAGGAGTTGCGACTGAAGGAGACCTTGAATAATCAGGGTGTCTCCAAATGCAGTGA 2086
Qy 1981 TCTGCTAGGCTGGTTCTCAGCTGCTCGGGCCCTTCTGCTCAGCGCGGGCTGCTGG 2040
Db 2087 TTGCTGTACCTGGTTCTCGGACAGCTTCTCATGATGATGAGCATCTTCTCTCTGA 2146
Qy 2041 TTCTGCTGCTCAAGCTGGGGACATCTCCCTTACAGCCACCCGGCGGTGTTCTCTGT 2100
Db 2147 CGATATTCATGATGATGAAGATTCCTACATTCAGCGACCCATTCATCTCTTCTGT 2206
Qy 2101 TCTTGGCAGCTTCCGCTGGCCAGCTGACCCAGAGCTTCTGCTCAGCGCTCTTCTCT 2160
Db 2207 TCTTGTGGCTTCTCCACTGCTGCCATCATCTGTGTCTTCTGCTCAGCACTTCTTCT 2266
Qy 2161 CCGCGCCAACTGCTGCGGCTCGGGCGGCTCGGCTTCTCTCTTCTTCTTCTTCTTCT 2220
Db 2267 CCAAGCCAGTCTGCGAGCAGCTGTAGTGTGTCTATCTATTTACCTCTTCTTCTTCT 2326
Qy 2221 ACGTCTGTGTGCTTGGGGGACCGCTGCTCCCGGGTGGCGGCTGGCGCGAGCC 2280
Db 2327 ACATCTGTGTCTGCTGGAGGACCGCATGACCGGTGAGCTGAAGAAGGCTGTGAGCT 2386
Qy 2281 TGCTGTCCCGCTGCTTCCGCTTCCGCTGCGAGAGCTTGTCTGTGAGGAGCAGG 2340
Db 2387 TACTGTCTCCGTTGCAATTGGATTGGCACTGAGTACTTGGTTGCTTTGAGAGCAAG 2446
Qy 2341 GGGAGCGCGCAGTGGGCAACAGTGGGACCCCGGCTTACGGCA---GACGTCTTACGCC 2397
Db 2447 GCTTGGGCTGCACTGGAGCAACATCGGGAAACAGTCCCAAGAGGGACGAATTCAGCT 2506
Qy 2398 TGGCCAGGTCTCTGGCTTCTGCTGTGAGCGCGGCTCTACGGCTCGCCACCTGT 2457
Db 2507 TCTCTCTGTCCATGACAGATGATGCTCTTGTGCTGTCTATGGCTTACTCGCTTGT 2566
Qy 2458 ACCTGGAAGCTGTGTGCCAGGCCAGTACGGGATCCCTGMAACATGGAATTTCTTTTC 2517
Db 2567 ACCTTGATCAGGTGTTTCCAGGAGACTATGNAACCCACTTCTTGGTACTTCTTCTAC 2626
Qy 2518 GGAGAGACTA-----CTGGTGC 2534
Db 2627 AAGAGTCGTATGTGCTTGGCGTGAAGGTGTTCACACAGAGAAAGAGCCCTGGAAA 2686
Qy 2535 GGACTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCACCCGCTGGACCCAAAGTG--- 2591
Db 2687 AGACGAGCCCTTAAACAGAGAAACGGAGGATCCAGAGACCCCAAGAGGAATACAGACT 2746
Qy 2592 --CTGGTGAAGAGGACACCGCGCTGAGTCTCTGGGCTATCCGTTCCGAGCTGGAGA 2649
Db 2747 CCTCTTTGAAGTGAAGATCAAGGGTGGTTCTTGGGGTATGCGTGAAGAAATCTGGTAA 2806
Qy 2650 AGCGCTTTCTGGAAGCCCGCAGCGAGCCCTGCGGGGGCTCAGCCTGGACTTCTTACCAAG 2709
Db 2807 AGATTTTGGAGCCCTGTGCGCGGCGAGCTGTGACCGTCTGAACATCACCTTCTTACGAGA 2866
Qy 2710 GCCACATCAGCGCTTCTTGGGCAACAGGGGCGGGCAAGACACACCTGCTCCATCT 2769
Db 2867 ACCAGATCACCGCATTTCTTGGGCCCAATATGAGCTGGGAAACCAACCACTTGTCCATCC 2926

Qy 2770 TGAGTGGCTCTTCCACCAGTGGTGGCTCTGCCTTATCCTGGGCCACGACGTCGCT 2829
Db 2927 TGAGGGTCTGTGTCACCAACCTCTGGGACTGTCTGTTGGGGAGGACATTGAAA 2986
Qy 2830 CCAGATGGCGCCATCCGGCCCCACCTGGGGGTCTGTCTCAGTACAAAGTGTGTTG 2889
Db 2987 CCAGCTGGATGCACTCCGGCAGAGCTTGGCATGTGTCCACAGACAAATCTCTGTTCC 3046
Qy 2890 ACATGCTCACCGTGAGCAGCAGCTGTGTTCTATGGCGGTGAAGGTCTGAGTCCG 2949
Db 3047 ACCACTCAGGTGCTGAGCAGATGCTGTCTATGCCAGCTGAAGAAAGTCCCAGG 3106
Qy 2950 CTGTAGTGGGCCCGGAGCAGGACCGTGTCTGCTGACGATGTGGGGTGGTCTTCAAAGCAGA 3009
Db 3107 AGGAGGCCAGCTGAGATGGAAGCAATGTTGGAGGACACAGGCTCTCACACAAAGCGA 3166
Qy 3010 GTGTGACAGTCCGCCACTCTCTGTTGGGATGCAACGGAAGTGTCTGCGTGGCATTTGCT 3069
Db 3167 ATGAAGAGGCTCAGGACCTATCAGTGGCATGTCAGAGAAAGTGTCTGCTTGCATTTGCT 3226
Qy 3070 TTGTGGCGGCTCCCAAGTGTGTTATCTTGGACGAGCTACGCTGGCGTGGATCTGCTT 3129
Db 3227 TTGTGGGAGATGCCAAGTGTGATTTGAGACGAACCCACCTCTGGGGTGGACCTTACT 3286
Qy 3130 CCGCGCGGTATTTGGGAGCTGTCTGCTCAAAATACGAGAAAGTGCACGCTGATCTCT 3189
Db 3287 CGAGAGCTCAATCTGGGATCTGTCTCTGAGTATCGCTCAGGCAGAACCATCATCTGT 3346
Qy 3190 CCACCCACCACTGGATGAGGAGAGTGTCTGGAGACCGTGTGGCTGTGTGGGAGGTG 3249
Db 3347 CCACTCACCATGAGCAGGCGGACCTCTTGGGACCGCATTTGCCATCATTTGCCCAGG 3406
Qy 3250 GCGCTTGTGCTGTGGCTCCCACTCTTCTGCGCGGTCACTGGGCTCCGCTACT 3309
Db 3407 GAAGCTCTACTGTCTAGGCACCCCACTCTCTCTGAAGAACTGTCTTGGCACAGGCTGT 3466
Qy 3310 ACCTGACCTGGTG-----AAGGCGCGCTCTCCCTGACCAACCAATGAGAAG 3357
Db 3467 ACTTAACTTGTGCGCAAGATGAATAACATCCAGAGCCAAAGGAAGGAGTGAAGGA 3526
Qy 3358 CTGACACTGAATGAGAGGCGAGTGTGGACACAGG-----CAGGAAAAAGAGA 3405
Db 3527 CCGTCACTGTCTGTAAAGGTTTCTCCACAGTGTCCAGCCCACTGATGATCTAA 3586
Qy 3406 ATGCGAGCCAGGCGAGAGTCTGGCACTCTCTAGCTGTCTGGCCCTGTGACAGACTGG 3465
Db 3587 CTCAGAACAAAGTCTGTGATGGGATGTAAATGAGCTGATGATGTCTTCCACCATG 3646
Qy 3466 TGCCCGGGGCACTGTGTGGAGGAGCTGCGACACAGAGCTGTGTGTGCTGCTGCCCTACA 3525
Db 3647 TTCCAGAGGCAAGCTGTGTGAGTGCAATTGCTCAAGAACTTATCTTCTTCTTCCAAATA 3706
Qy 3526 CGGTGCGCATGACGGCAGCTTCCCACTCTTTCGAGAGCTAGACACGGGCTGGCGG 3585
Db 3707 AGAACTTCAAGCACAGAGCATATCCAGGCTTTTTCAGAGAGCTGGAGAGACGCTGGCTG 3766
Qy 3586 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAGGAGATCTTCTGAGG 3645
Db 3767 ACCTTGTCTCAGCAGTTTGGAAATTTCTGACATCTCCCTCGGAAGAGATTTTCTGAGG 3826
Qy 3646 TGTGTGAGGAGTGTGTGCGGACACAGATATGAGGATGGG-----AGCTGCGGGCAGC 3699
Db 3827 TCACGGAGGATTTGATTTTCAAGACTCTGTGTTGGGGTGGGCTCAGCAGAAAGAGAAA 3886
Qy 3700 ACCTATGACAGGCAATTTGCTGGCTTAGACGTAACTCTGCGGCTCAAGATGCGGCCACAGG 3759
Db 3887 ACGTCAACCCCGGACACCCCTGCTGGTCCAGAGAGAGGCTGGACAGACACCCCGAGG 3946
Qy 3760 AGACAGGCTGGAGAACGGGGAAACAGTGGGTGAGCCCCAGAGACTGACACAGGGCTCTG 3819
Db 3947 ACTCAATGTCTGCTCCCAAGGGCGCGGCTGTCTACCCAGAGGGCGGCTCTCCCCAG 4006

Db 6080 ACTGCTGCTGGCAGTACAGTGGGGGCAACAAGGGAACTCTCCACAGCCATCGCAC 6139
Qy 6028 TGGTTGGGACACAGCCGTGGTGTCTTGGAGAGCGGACGACGATGACCCACAGG 6087
Db 6140 TCATTGGCTGCCACCGCTGGTGTCTGGATGAGCCACACAGGATGACCCACAGG 6199
Qy 6088 CCGCGGCTTCTTTGGAAACAGCTTTTGGCGGTGGTGGGAGGGCGTTTCAAGTATGC 6147
Db 6200 CACGCGCATGCTGTGGAACGTCATCGTAGCATCATCAGAGAAGGAGGGCTGTGGTCC 6259
Qy 6148 TCACCTCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGGCGCTAGCCATCATGTGA 6207
Db 6260 TCACATCCACAGCATGAGAGATGTGAGCACTGTGTACCGGCTGCCCATCATGGTAA 6319
Qy 6208 ATGGCGGTTCCGCTCGGCGGACCCGCGCAACATCTCAAGGGCAGATTCGCGGGGTC 6267
Db 6320 AGGGCGCTTTTCATGTATGGGACCATTCAGCATCTCAAGTCCAAATTTGGAGATGCT 6379
Qy 6268 ACACACTGACCTGGGGTGGCGCGCAAGTCCAGCGGCGAGCGCTTCGTGGGG 6327
Db 6380 ATATCGTCACAAATGAAGATCAAAATCCCGAAGGAGCAGCTGCTTCTGACCTGAACCTG 6439
Qy 6328 CGAG-----TTCCCTCGGTGGAGTGGCGAGGACATGGAGGCGCGCTGCGCTTCC 6381
Db 6440 TGGAGCAGTTCTTCAGGGGAATCTCCAGGAGTGTGAGGGAGGAGGACATCAACA 6499
Qy 6382 AGCTGCGG---CCGGGAGGGGCTGCGGCCCTGGCGCGGTCTTTGGAGAGCTGGCGTGC 6438
Db 6500 TGCTCCAGTTCAGGTCC 6559
Qy 6439 ACGGCGCAGACGCGGTGAGGACTTTTCGTGAGCCAGACAGATGCTGAGAGGAT 6498
Db 6560 ACAAGGACAGCGCTGCTCATCGAGGAGTACTCAGTCACACAGACACACTGGACCGGTGT 6619
Qy 6499 TCTGTACTTCTCAAGGACGAG 6521
Db 6620 TTGTAAATTTTGTAAACAGCAG 6642

RESULT 5

US-09-774-528-209
; Sequence 209, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aiqiong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: Pt_Fl_genes Version 2.0
; SEQ ID NO 209
; LENGTH: 8037
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1654) .. (7329)
US-09-774-528-209

Query Match 13.5%; Score 920; DB 4; Length 8037;
Best Local Similarity 53.3%; Pred. No. 5.2e-183;
Matches 2832; Conservative 0; Mismatches 2125; Indels 357; Gaps 28;
Qy 1593 CCCGGCCACGTGGCGCATCAAAATCCGATCGACATTTGAGTGTGTCAGAGGACCAATAA 1642
Db 1854 CCCGCTCACGTGCACACTACAAGATCCGCAAGACTCCAGCTTTCACCGAGAAACAACGA 1913
Qy 1643 GATCAGGAGCAGGTTTGGGACCTTGGCCAGCGCGGACCCCTGAGCCGACCTCGGCTA 1702
Db 1914 GATCGCGCGCTTACTGGCGGCTTGGCCCAATACTGCGCGCGCTTCTACTCTCTA 1973
Qy 1703 CGTGTGGGCGGCTTCTGTGTACTCTGCAAGACTTGGTGGAGGTGCGAGCGCTCGGCTCT 1762
Db 1974 -----CGGCTTGGTCTGGATCCAGGACATGATGAGCGCGCATCATCGACACTT 2024
Qy 1763 CAGCGGCGCAACCCCGGG---CGGCTCTACTGCGAGCAGATGCCCTATCCGTGCTA 1819
Db 2025 TGTGGGCGACACGTGGTGGAGCCAGGCGAGCTAGTGCAGATGTTCCCTTACCCTGCTA 2084
Qy 1820 TGTGAGCAGCGTGTCTGCGTGTGCTGAGCGGCTCGCTGCGGCTTCTTCTGAGCTGGC 1879
Db 2085 CACAGCGATGACTTCTGTTTGTATTTGATGAGCAGATGATGCCCTGTGCTGATGCTC 2144
Qy 1880 CTGATCTTCTCTGACACTGACAGTGAAGCGCGTGGTGGCGGAGAGAGAGCGCGCT 1939
Db 2145 CTGGGTCTTCTCTGCGCATGACCATCCAGCAGATCGTGGCGGAGAGAGAGCGCGCT 2204
Qy 1940 GCGGAGACACATGCGCGCATGGGCTCAGCGCGCGGTGTCTGGCTAGGCTGGTCTCT 1999
Db 2205 CAAGGAGTGTGAAGACCATGGGCTGAAACAACGCGGTGCACTGGGTGGCTGGTCTCAT 2264
Qy 2000 CAGTGTCTGGGCGCTTCTGCTCAGCGCGGCTGCTGCTTCTGCTGCTCAAGCTGGG 2059
Db 2265 CACCGCTTGTGAGCTGTCCATCTCGTGAAGAGCTCAACGCGATCTCTGAAGTACGG 2324
Qy 2060 GGACATCTCTCCCTACAGCCACCGCGGCGTGTCTTCTGTTTGGCAGCGCTTCCGCGT 2119
Db 2325 CCAGTGTCTATGACAGCCAGCGGTGTCATCATCTGGCTCTTCTGGCAGCTTACGCGT 2384
Qy 2120 GGCACGCGTACACAGAGTTCCTGCTCAGCGCTTCTTCTCCGCGGCAACCTGGCTGC 2179
Db 2385 GGCACCATCATGTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2444
Qy 2180 GGCCTGGGCGGCGCTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2227
Db 2445 GGCCTGGGCGGCGCATCATCT 2504
Qy 2228 GTGTGTGCTTGGCGGACCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2287
Db 2505 AGAGAGGTGGCGCATGATAGATCACGGCTTCGAGAGTGCATCGGCTCCCTCATGTC 2564
Qy 2288 GCGCGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2347
Db 2565 CACGACGCGCTTGGTCTGGGCTCTAAGTACTTCTGCGCTGTATGAGTGGCGGCGGCTGG 2624
Qy 2348 CGCGCAGTGGCACAACCTGGGCA---CCGCGCTTACCGCAGACCTTCTCAGCTGGGCCA 2404
Db 2625 CATCAGTGGCAGACCTTTCAGCCAGTCCCGGCTGGAGGGGAGGACTTCAACTTGTCTCT 2684
Qy 2405 GGTCTCTGGCTTCTGCTGTGAGCGGCGGCTCTAGGCGCTCGCCACTCGCTTACCTGGA 2464
Db 2685 GGTGTACCATGCTGATGGTGGACCGCTGGTCTATGGCATCTCTCAGTGGTACATGA 2744
Qy 2465 AGCTGTGCGGCGGCGGAGTCCCTGAAACCATGGAATTTTCTTTTTCGAGAGAG 2524
Db 2745 GGTGTGCAACCGGCGGCTGACGCGGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2804
Qy 2525 CTACTG-----GTGCGGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2574
Db 2805 CTACTGGCTGGGAGTGGCGGCGGAGAGAGCTTGGAGTGGAGCTGGCGGCGGCGGCGGCGG 2864

Qy	2575	CGCTGGACCCAAAGGTCGCTGG-----	2599
Db	2865	CCCCCGCCTCAGTGTCTATGGAGGAGGACACAGGCCTGTGCCATGGAGAGCGGGCTTTTGA	2924
Qy	2596	-----TAGAAGGACACCGCCGGCCTGAGTCTCTGGCGTATCCGTTGCGCAG	2641
Db	2925	GGAGACCCGTGGCATATGAGGAGGAGCCACCCACCTGCCTTGCTGTCTGCGTGGACAA	2984
Qy	2642	CCTGGGAAGCGCTTCTCTGGAAGCCCGCAGCAGCCCTCTGGGGGCTCAGCCTGAGACT	2701
Db	2985	ACTCACCAAGGTTACAAAGGACGACAGAGCTGGCCCTGAAACAAAGCTGAGCCTGAACCT	3044
Qy	2702	CTACGAGGCCACATACCCGCTTCTTGGGCCAACACGGGGCGGGCAAGACCAACCAACCT	2761
Db	3045	CTACGAAACCAAGTGTCTCTCTTCTTGGGCCAACACGGGGCGGGCAAGACCAACCAACCT	3104
Qy	2762	GTCCATCTTGAGTGGCCTCTTCCACCCAGTGTGGCTCTGCCCTTCACTCTGGGCCACGA	2821
Db	3105	GTCCATCTTACCGGCGCTGTTCCTCCAAAGCTCGGGTTCCGCCACCACTATACCGGGCACGA	3164
Qy	2822	CGTCGCGCTCAGCATATGGCCGCATCTCCGGCCCCACCTGGGCGTGTGCTCCTCAGTACAACT	2881
Db	3165	CATCCGACCGAGATGATGAGATCCCGAAGACCTGGGCAATGTGCCCGCAGACATATGT	3224
Qy	2882	GCTGTTTGATGCTGACCGGTGGACGACGCTGTGTTCTATGGGCGGCTGAAGGCTCT	2941
Db	3225	GCTCTTTGACCGGCTCACGGTGGAGGAACCTCTCTGTTCTACTCACGGCTCAAGAGCAT	3284
Qy	2942	GAGTGCCTGTAGTGGGCCCCGAGCAGGACCTGCTGCTGAGGATGTGGGCTGTGCTC	3001
Db	3285	GGCTCAGAGGAGATCCGACAGAGATGGAACAAGATGATCGAGACCTGGAGCTCTCCAA	3344
Qy	3002	CAAGCAGAGTGCAGACTCGCCACTCTCTGGTGGATGCAACGGAAGCTGTCCGTGGC	3061
Db	3345	CAAAAGCACTACTGTTGACAGATTTGTGCGGTGGCATGAGCGCAAGCTGTCCGTGGC	3404
Qy	3062	CATTGCCCTTTTGGGCGGCTCCCAAGTTGTTATCTGGAGAGCCTACGGCTGGCGTGA	3121
Db	3405	CATGCGCTTCGTGGCGGCTCTCGCGCCATCATCTCGAGAGCCACGCGCGGCGTGA	3464
Qy	3122	TCCTGCTTCCGCGCGGATTTTGGAGCTGTGCTCAAAATACGGAAGTGTGACGCT	3181
Db	3465	CCCCATACGCGCGCGCCATCTGGGACCTCATCTGGAAGTACAAGCAGCGCGCGCACCAT	3524
Qy	3182	GATCCTCTCACCCACCACTGGATGAGGACAGCTGCTCGGAGACCGTGTGCTGTGGT	3241
Db	3525	CGTTCTGTCAACCCACACATGATGAGGCTGACCTGCTTGGGACCGCATATGCCATCAT	3584
Qy	3242	GGCAGTGGCGCTGTGTGCTGTGCTGCCCTCCCACTCTCTGCGCGCTCACCTGGGCTC	3301
Db	3585	CTCCATGGGAAGCTCAAGTGTCTGCGCTCCCGCTCTTCTCAAGGGACCTATGGCGA	3644
Qy	3302	CGGCTACTACTGACGCTGTGTAAGGCCCGCGCTGCCCTGACCAACCAATGAGAAGGCTGA	3361
Db	3645	CGGGTACCGCTCACTACGCTGTGTAAGCGGCCCGCGAGCGGGGGCCCCCAAGAGCCAGG	3704
Qy	3362	CACTGACATGGAGGGCAGTGTGGACACAGGACGGAAGAAAGNATGGCAGCGCAGGGCAG	3421
Db	3705	GCTGGCATTCAGCCCCCAGGTGGGGCCCCGCGCTGAGCAGTGTCTCCAGTCTCAGGTGTC	3764
Qy	3422	CAGAGTCGGCACTCCTCAGCTGTCTGGCCCTGGTACACACTGGGTGTCGCGGGGCACGGCT	3481
Db	3765	CCAGTTCTTCGCAAGCATGTGGCTCTCTGCCCTGCTGTCTCAGACAAAGCAGCGAGCT	3824
Qy	3482	GGTGGAGGAGCTGCCACACGAGCTGGTG-----CTGGTGTGTGCC	3520
Db	3825	CTCCTACATCTGTGCCACGAGCGCGCCAAAGAGGGGGCTTTTCAGCGCCTCTTTCAGCA	3884
Qy	3521	CTACACGGTGCCTATACGGCAGCTTGGCCACACTCTTCCGAGAGCTACACCGCGCT	3580
Db	3885	CCTGGAGCGCAGCCTGTGATGCATGTCACCTCAGCAGCTTTTGGGCTGTATGACACGACCT	3944
Qy	3581	GGCGGA-----GCTGAGGCTCACTGGCTACGG	3607

Db	3945	GSAGGAAGTGTTCCTCAAGGTGTCGAGGAGGATCAGTGCCTGGAGAACAGTGAGGCCGA	4004
Qy	3608	GATCTCCGACACACAGCCTCGAGGAGATCTTCTCGAAGGTGGTGGAGAGTGTCTGCGGA	3667
Db	4005	TGTGAAGGAGTCCAGGAGATGTGTCTCCTGGGGCGAGGGCCCGCGTCTCGGGAGGG	4064
Qy	3668	CACAGATATGAGGATGGCAGTGCGGGAGC--ACCTATGCACAGCAATGCTGAGCCTA	3725
Db	4065	TCACGCTGGCAATCTGGCCCGTGTCTGGAGCTGACCCAGTTCGACGATCGCTGCAGTC	4124
Qy	3726	GAGGTAAACCTCGGGCTCAAGATCGCGCCACAGGACACAGCGT-----	3769
Db	4125	GGCGTCACTGTGGGCTCTGCCGTGGCGACGAGGAGCTGGCTACACCGACGTCTATGG	4184
Qy	3770	GGAGACGGGGAAACAGCTGGGTGAGCCCCAGAGACTGACACGGGTCTTGGGCCAGACGC	3829
Db	4185	CGACTACCGCCCTCTTTGATAACCCACAGAACCCAGACAATGTGAGCTGCAAGGT	4244
Qy	3830	CGTGG-----GCGGGTACAGGGCTGGGCACCT	3856
Db	4245	GGAGGCAGAGGCCCTGTGAGGGTTCGCCACGGGACGCCGAAGCTGGAACGGGTGGCT	4304
Qy	3957	GACCCGCCAGCAGTTCAGGCCCTGCTTCTCAAGCGGTTTTCTGCTTGGCCGCCGACGCG	3916
Db	4305	GAAGGTGCGCAGTTCACCGGGTGTGTGTCAAACGCTTCACCTGGCGCCCGCGCACTC	4364
Qy	3917	CCGGCGCTGTTCGCCCCAGATCGTGTGCTGCTGCCCTTTGTGGGCTGSGCCCTCGTGT	3976
Db	4365	CAAGGCACTCTTCTCCAGATCTTGTCTGCAGCCCTTCTGCTGTGGTGCCATGACCGT	4424
Qy	3977	CAGCCTCATCGTGCCTCTTTCCGGGCACTACCCGGTCTGCGGCTCAGTCCCAACCATGTA	4036
Db	4425	GGCCCTGTCGTCCCGGAGATTGGTATCTGCCCCCGCTGCTCTGTCACTTCCAGTA	4484
Qy	4037	C-----GGTGTCAAGTGTCTTCTTCAAGTAGAGACGCCCCAGG	4075
Db	4485	CCACAACCTACACCCAGCCCGTGGCAATTCATCCCTACGCCAACGAGAGCGCGCGA	4544
Qy	4076	GGACC-----CTGGAAGTGCCTGGCTCTCGAGCGCTGTGCA	4114
Db	4545	GTACCGCTGCGGTATGCGCCGACGCGACGCCCCAGAGCTCGTAGACAGTTCGGCT	4504
Qy	4115	GGAGGCAGGACTGGAGAGCCCCAGTG--CAGCATAGCTCCACAGGTTCTCGGCACCA	4172
Db	4605	GCCGTGGGGTGGTGGTCCACCTGCGTGTCAAGTCTCCGCCAACGGCTCGTGGGCC	4664
Qy	4173	GAAATTCTGCTGAAGTGGCAAGGTCTTGGCCAGTGGCAACTGG-----ACCCAGAGT	4227
Db	4665	CAGTTGAACCTGACAGACGGGAGTGGCCCTGCTGGCGGCTCGGTTCTTCACAGCAT	4724
Qy	4228	CTCCATCCCGAGCTGCCAGTGTAGCAGCCCGGTGCGCGGCTGCTGCCCGAGTGC	4287
Db	4725	GTGTCTGGAGTCTTACACAGGGGTGCGACATGTCCAAATTTCTGTCACCCGCCCTC	4784
Qy	4288	CGGCTGCAGTGTGTGTCCTCCCTCCGCCCCAGGACGTACCGGCTCTGGGGAAGTGGTTC	4347
Db	4785	GCCGCCCACTGATCTGCGCAGGTTCCCGGATGAGAACCTGAGGCCTGGAACTCTC	4844
Qy	4348	AGAACTTGACAGGCCGGAACTGTCTGACTTCC--TGGTCAAGACCTTACCCGCGCTGGT	4405
Db	4845	CCTGCCCCCAACGCTGGGCGCAGAAATGTGGACGTGGCACCCCTCCCTGTCGCGGCTGGT	4904
Qy	4406	GCGCCAGGCCTGAAGACTAAGAGTGGGTGATGAGTCAAGTACGAGGCTTCTCGCT	4465
Db	4905	ACGGGAGCCCGTCCGTGCACTGCTGTGCGCAGGGACACCGGCTTCTCTGCCCCAGCAG	4964
Qy	4466	GGGGGGCCGAGACCCAGGCCCTG-----CCC	4490
Db	4965	TGTGGGCGGCAACCGCCCAAGATGCGGGTGGTTCAGGGGACATCCTGACCCGACATCAC	5024
Qy	4491	TCGGGCCAAGATTGGGCGGCTCAGTGGAGGAGTTGTGGGCGTGTGAGTCCCTGCCT	4550

Db 5025 CGGCCACATGCTCTCTGAGTACCTGCTCTTACCTCCGACCGCTTCCGAGTCGACCGGTA 5084
 Qy 4551 GCGCG---GGCCCTCGACCGGTCTCTGAAACCTCTACAGCTTGGGCTCTACAGCTTGA 4606
 Db 5085 TGGGCCCATACCTTTGGAAAGCTCTGAAAGTCCATCCAGCTCATTTGGCACCAGGSC 5144
 Qy 4607 TGCTCAGACAGTCTCAAGATCTGTT-----CAACAA 4642
 Db 5145 CCCACCCATGTCGCGAAGATCGCGGTGCGCAGGGCTGCCAGGTTTCTTACAAACA 5204
 Qy 4643 AGGCTGGCACTCCATGTTGGCTTTGCAACCGAGCCAGCAACCAATCTCCGTCCTCA 4702
 Db 5205 GGGCTATCAGACATGCCCACCTACTCAACAGCTCAACAGCGCATCTCTGCGGCCAA 5264
 Qy 4703 CTTGCCCCAGGCGCGCCGCGC---AGCCCAACAGCATCAACACCTCAACACCCCTT 4759
 Db 5265 CTTGCCCAAGAGCAAGGCAACCCGCGGCTTACGGCATCCGTCACCAACACCCCAT 5324
 Qy 4760 GAACTCAGCAAGGAGAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGGTGAGCTCT 4819
 Db 5325 GAAATAGACCAAGCGCCAGCTCTCCCTGGATTTACCTGCTGAGGGCAC---GGATGTCGT 5381
 Qy 4820 GGTCTCCATCTGTGTGCTTTTGCCATGTCCTTTGTCGCGGCACTTCACTCTGTCTCT 4879
 Db 5382 CATGCCATCTTCACTCATCTGTGGCCATGTCCTTCTGTCGCGGCACTTCTGTCTCT 5441
 Qy 4880 CATTGAGGAGCGAGTCACCGAGCCAGCACTGAGCTCATGCGGGGCTGTCCCCAC 4939
 Db 5442 CGTGCGCGAGAGTCCACCAAGGCCAAGCATCTGAGTTGTACAGGCTGCAACCCCAT 5501
 Qy 4940 CTTCTACTGCTTGGCAACTTTCTCTGGGACAT-----GTGTAAC 4981
 Db 5502 CATCTACTGCTGCGGAACACTAGTGTGGGACATGCGCGCCCAACCCCTGACAGCTCAA 5561
 Qy 4982 CTTGTCGCGAGATGATGCTGCTGCTCATCTTTCTGCGCTTCCAGAGAGGATATGT 5041
 Db 5562 CTTGTCGCGGCTTACCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5621
 Qy 5042 GCGCCCTGCCAATCTGCT 5101
 Db 5622 GTGCGCCACCACTTCCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5681
 Qy 5102 ACCGCTCATGACCGAGCT 5161
 Db 5682 GCGCATCATGACCGGCT 5741
 Qy 5162 CACTGCAATAACCTTTATGAGATCAATGGAAGATGCGACCTTTGCTGAGCT 5221
 Db 5742 CATTTGTCATCAATCTCTTCTATCGGCATCACCAGCCACCGTGGCCACCTTCTCTCTCT 5801
 Qy 5222 CTTC---TCTGATCAGAGCTGACGAGGTGAGCGGATCTTTGAAACAGGCTTCTCTCTAT 5278
 Db 5802 CTTGAGCAGCAGCAGGACCTGAGGTTGTCAACAGTTACTTGAAGCTGCTTCTCTCTCT 5861
 Qy 5279 CTTCCCCCATCTCTTTGGGCGGGGCTCAITGACATGTTGCGGAACACGAGGCTGAGC 5338
 Db 5862 TTTCCCCCACTAGCAACTGTCGCGCCAGGCTCATGAGATGCGCTTACACAGAGTACATA 5921
 Qy 5339 TGATGCTTTGAGCGCTTGGAGACAGGC---AGTTCCAGTCAACCCCTGCGCTGGAGGT 5395
 Db 5922 CGAGTACTACGCCAAGATTTGCAAGATGAAAGTCCCGGTTTCGAGTGGGACAT 5981
 Qy 5396 GGTGGCAAGAACCTTTGGCCATGTTGATACAGGGGCCCTCTCTCTCTCTCTCTCTCTCT 5455
 Db 5982 TGTACCCCGGACTTGTGGCCATGGCGGTTGAGGGCGCTGTGGGTTTCTCTCTCTCTCT 6041
 Qy 5456 ACTGTCGACGACCGAAGCAACTCTCTGCCACAGCCAGGCTGAGGTCTCTGCACTCTCT 5515
 Db 6042 CATGTGCCAGTACAATCTCTCTGCGGCGCCACAGGCGATGCTGTGTCT---ACCAAGCC 6098
 Qy 5516 GGGAGAGGAGCAGAGGATGTAGCCCGGTGAACGGGAGCGGCTGTGCAAGAGCACCCA 5575
 Db 6099 TGTGGAGGATGTGAGCGTGGCGGCGAGTGTGAGCGGCGAGGAGTGTCTCCGGGAGACGCG 6158

RESULT 6
 US-08-259-24
 ; Sequence 24, Application US/08665259

Qy 5576 GGGGATGTGTTGGTGTGAGGAACTTGACCAAGGTATACCGTGGGCGA-----G 5626
 Db 6159 CAATGACATGTCAAGATTGAGAACTTGACCAAGGTCTTACAAGTCCCGAGATTGGCG 6218
 Qy 5627 GATGCGAGCTTTGACCGCTTGTGCTGGGATTTCCCTCTGGTGTGATGTTTGGCTGCT 5686
 Db 6219 TATCTGCGCTTGTGACCGCTTGTGCTGGGTGTGCTCTGCGAGTGTCTTGGGCTCTCT 6278
 Qy 5687 GGGTGTGAATGAGCAGGGAAGAGTCCACAGTTCGATGTCGAGGGGACACATTGGC 5746
 Db 6279 GGGGCTCAACGCTGCGGCAAGACACAGCACTTCAAGATGCTGACCGCGAGAGGAC 6338
 Qy 5747 CAGCAGGCGGAGGCTGTGCTGGCAGGCCACAGCGGTGCGCGGAACCCAGTGTCTGCGCA 5806
 Db 6339 GACGGGGCGGAGGCTTCTGTCATGAGACACAGCTGCTGAAGGAGCTGCTCCAGTGCA 6398
 Qy 5807 CTTGAGCATGGGATCTGCTCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCGGCA 5866
 Db 6399 GCAGAGCTCTGCTTACTGCGCGAGTGTGACGCGCTGTTTCGACGAGCTTCAAGGCGGGA 6458
 Qy 5867 GCACCTGAGCTGTGCGGCGCTGCGGCTGTCGAGGCGCCAGGTTGCCAGACGCG 5926
 Db 6459 GCACCTGAGCTGTACAGCGGCTGCGTGGGATCTCTGGAAGGACGAGGCGCGGCTGT 6518
 Qy 5927 TGGTCTCAGGCTGCGGCTCTGAGCTCTCTATGTTAGCAGACCGGCTGTCAGGCACTTA 5986
 Db 6519 GAAGTGGGCTCTGAGAGCTGAGGCTGACCAAGTACGACCAAGCGGCTGGCACCTTA 6578
 Qy 5987 CAGCGGAGGGAACAAACGAACTGGGAGCGGCTGCTGCTGCTGTTGGGACCCAGCGCT 6046
 Db 6579 CAGCGGCGGCAACAGGAGGAGCTCTCCAGCGCATGAGCTTCTGAGTGTGCTGCTGCTG 6638
 Qy 6047 GGTCTTCTGAGCAGCGGACCAAGGATGAGCCAGCGGCGGCTTCTCTTTGAA 6106
 Db 6639 CATCTTCTGAGCAGGCGGACCAAGGATGAGCCCAAGGCGCGGCTTCTCTTTGAA 6698
 Qy 6107 CAGCGCTTGTGCGGAGGCGGCTTCTGATGATGCTCACTCTCCATAGCATGGA 6166
 Db 6699 CTTCTCTGAGCTCATCAAGACAGGCGGTTCTGATGCTGATCATCACAGCATGGA 6758
 Qy 6167 GAGTGTGAAGCTCTGCTGCGCTTACCATCATGTTGATGAGGCTGCGGCTGCTGCT 6226
 Db 6759 GAGTGTGAGCGCTGTGACGCGCTGCGCTGCGCTCATGTTGAAAGCTGCGCTGCGT 6818
 Qy 6227 GGGCAGCGGCAACATCTCAGGCGGATTCGCGCGGCTCAGCATGACCTGCGG--- 6284
 Db 6819 GGGCAGCATCCAGCACCTGAAGAACCGGTTTGGAGATGGCTTACATGATCAGGTCG 6878
 Qy 6285 -GTGCGCGCGCAAGTCCAGCGCGGCTTCTGCGCGGCTTCTGCGGCTTCTGCGGCT 6343
 Db 6879 CAAGACGACGAGTGTGAAGAGCTGTTGCGGTTCTTCAACCGCAACTTCTCCGGAAGC 6938
 Qy 6344 GAGCTCGCGAGGACATGAGAGCGCTTCTGCTTCCAGCTGCGCGGAGGCGGCTG 6403
 Db 6939 CATGCTCAAGAGGCGGACCAACAAAGGTGCACTACAGCTCAAGTCG---GAGCAT 6995
 Qy 6404 CGCCCTGGCGGCTTCTTGGAGAGCTGGCGGTCACGCGGCGAGACGCGGCTGAGGA 6463
 Db 6996 CTGCTGCGCGGCTTCTGAGCAAGATGAGAGCTGCTGCGGCTGCTGCGGCTGAGGA 7055
 Qy 6464 CTTTCTGCTGAGCAGCATGCTGAGGAGGATTTCTTGTACTTCTTCAAGGACCGAGG 6523
 Db 7056 CTACTCGCTGAGCAGGACCACTGAGACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 7115
 Qy 6524 GAAGGACGAGGACCAACGAGAGCAGAGGAGCA 6557
 Db 7116 TGACAACTGGAGCAGGAGACGAGCGGCGCA 7149


```

; Patent No. 6028173
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..5053
; US-08-665-259-24

Query Match          5.9%; Score 404.6; DB 3; Length 5894;
Best Local Similarity 55.8%; Pred. No. 3.3e-75;
Matches 876; Conservative 0; Mismatches 659; Indels 36; Gaps 4;

Qy      1793  CCTGCAGCAGATGCCCTATCCGTGCTATGTGGAGCAGGTGTTCTGGTGGTGTGCTGAGCCG 1852
Db      664  CATCAAGAGGTTCCCGTACC CGCGGTTTCATCGCAGACCCCTTCTCTGTGGCCATCCAGTA 723

Qy      1853  GTCCGTCGCCGCTCTTCTTCGTACGGCTGGCGCTGGATCTACTCCGTGACACTGACAGTGAAGGC 1912
Db      724  CCAGCTGCCCTGCTGCTGCTGCTGCTCAGCTTCACTACACCGGCTCACCATTGCCCGTGC 783

Qy      1913  CTGTGTCGGGAGAGAGACGCGGCTGCGGACACCATCGGCGCCATNGGGGCTCAGCCG 1972
Db      784  TGTGTCGAGGAGAAAGAAAGAGGCGCTGAAGGAGTACATCGCATGATGGGGCTCAGCAG 843

Qy      1973  CGCGGTGCTCTGGCTAGCGTGGTTCCTCAGTCGCTCGGGCCCTTCTGCTCAGCGCCGC 2032
Db      844  CTGGCTGCATGGAGTGCCTGGTTCCTTGTGTTCTTCTCTTCTCTCATCGCGCGCTC 903

Qy      2033  GCTGCTGGTTCGTGTGCTCAAGCT-----GGGGGACATCCTCCCTACAG 2077
Db      904  CTTTCATGACCTGCTCTTCTGTGTGTCAGGTGAAGCCAAATGTAGCGGTGCTGTCCCGCAG 963

Qy      2078  CCACCCGGGGCGTGGTCTTCTCTGTTCTTTGAGGACGCTTCGCGGGTGCCACGGTGACCCACAG 2137
Db      964  CGACCCCTCCCTGGTGTGCTGCTTCTGCTGTGTGTCGCCATCTCTACCATCTCTTTCAG 1023

```

QY 3197 CCACCTGATGAGGAGAGCTGCTGGAGACCGTGTGGCTGTGGTGGAGGTGGCGGCTT 3256
DB 2104 CTTTCATGACGAGGTGTGACCTGCTGGAGACCGCATCGCCATCATGAGGAGGAGCT 2163
QY 3257 GTGCTGTGTGCTCCCACTCTTCTCGCGCGTCACTGGGCTCCGGCTACTACTGAC 3316
DB 2164 GCAGTGTGTGGGTCTGCTGCTTCTCAAGCAGAAATACGGTGGCGGCTATCATGAC 2223
QY 3317 GCTGTGAAGG 3327
DB 2224 GCTGTGAAGG 2234

RESULT 7

US-08-762-500-24
; Sequence 24, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..5053
US-08-762-500-24

Query Match 5.9%; Score 404.6; DB 3; Length 5894;
Best Local Similarity 55.8%; Pred. No. 3.3e-75;
Matches 876; Conservative 0; Mismatches 659; Indels 36; Gaps 4;
QY 1793 CCTGACGATGCCCTATCCGTCGTATGTGGACGACGTGTTCCTGGCTGTGCTGAGCGG 1852

DB 664 CATCAAGAGGTTCCTGATACCCGCGCTTCATCGCAGACCCCTTCCTCGTGCCATCCAGTA 723
QY 1853 GTGCTGCGGCTCTTCTTGACGCTGGCTGATCTACTCGGTGACACTGACAGTGAAGGC 1912
DB 724 CCAGTGGCCCTGTGTGCTGTCTAGCTTCACTACACCGGCTCACCATTGCCCGTGC 783
QY 1913 CGTGTGGGAGAGAGACGCGCTGCGGGACACCATGCGCGCATGGGGCTCAGCGG 1972
DB 784 TGTGTGAGGAG 843
QY 1973 CGCGGTGCTCTGGCTAGGCTGCTTCTCAGCTGCTCGGGCCCTTCTGCTCAGGGCGGC 2032
DB 844 CTGGCTGCACGTGAGTGCCTGGTTCCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 903
QY 2033 GCTGCTGCTTCTGCTCAAGCT-----GGGGACATCTCTCCCTACAG 2077
DB 904 CTTTCATGACCTGTCTTCTGTGTCAAGGTGAAGCCAAATGTAGCCGTGTGTCTCCGAG 963
QY 2078 CCACCCGGGGGTGTCTTCTGTTCTTGGAGCCTTCGCGGTGGCCACGCTGACCCAGAG 2137
DB 964 CGACCCCTCCCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
QY 2138 CTTCTGCTCAGCGCTTCTTCTCCGGGCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2197
DB 1024 CTTTCATGTCAGCACCTTCTTTCAGCAAGCCAAACATGCGCAGCAGCTTTCGAGGCTTCT 1083
QY 2198 CTACTTCTCCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2257
DB 1084 CTACTTCTTACCTACATCT 1143
QY 2258 GGGTGGCGGCTGCGCGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2317
DB 1144 GAGCCAGAGCTCTGCT 1203
QY 2318 CTTGCTGCTGTGAGAGAGAGGCGAGGCGCGCGCAGTGGC---ACAACTGGGCACTCG 2374
DB 1204 CATTTGGGAAATTTGAGGCGAAAGCATGGGCATCCAGTGGCGAGACCTCTCTGAGTCC 1263
QY 2375 GCCTACGCGAGAGCTTTCAGCCTGGCCAGGCTCTCTGCGCTTCTGCTGCTGCTGCTG 2434
DB 1264 CAAGTGGACAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1323
QY 2435 GCTCTACGGCTCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2494
DB 1324 GCTCTATGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383
QY 2495 TGAACATGGAATTTTCT 2548
DB 1384 TCAGCCTGTGTCTTCT 1443
QY 2549 -----CAAGAGTCCAGCCCTTGGCCCGCCCGCTGAGCCCAAGGTGCTGTAGAGA 2602
DB 1444 AGGGAAGGAG 1503
QY 2603 GGCAACCCCGGCTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2662
DB 1504 CGAGCCAGAGAGCTGTGTGGGGGATCAAGATCAAGCACTGTCTCAAGGTGTTTCAAGGT 1563
QY 2663 AAGCCCGCA-----GCCAGCCCTGCGGGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTG 2716
DB 1564 GGGAAATAAGGACAGGGCGCGCTCAGAGACCTCAACCTCAACCTGATACAGGAGACAGAT 1623
QY 2717 CACGCTTCTCTGGGCGACAAAGGGGCGGCGAGACCAACCACTGCTGCTGCTGCTGCTGCT 2776
DB 1624 CACGCTTCTGTGGGCGACAAAGGGGCGGCGAGACCAACCACTGCTGCTGCTGCTGCTG 1683
QY 2777 CCTCTTCCACCCAGTGTGCTCTGCTTCTATCTCTGGGCGACAGACCTGCTGCTGCTGCT 2836
DB 1684 TCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1743
QY 2837 GGCGCGCATCCGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2896

1744	Db	GGTTTCAGATCCGGNAGAGCCTGGCCCTGTGCCCCACACACGACATCTCTGTTTGACACTT	1803
2897	Qy	GACCGTGGACGAGCAGCTCTGGTTCTATATGGCGCGCTGAAGGCTGTAGTGCCCGCTGTAGT	2956
1804	Db	GACAGTCGCAGAGCACCCTTTATTTCTACGCCACAGCTGAAGGGCTGTCACTCAGAAAGTG	1863
2957	Qy	GGGCCCCGACGACGAGACCGTCTGCTGCAGGATGTGGGGCTGGTCTCCAAAGCAGAGTGTGCA	3016
1864	Db	CCCTGAAGAAGTCAAGCAGATGCTGCAATCATCGGCTTGAGACAAAGTGGAACTCAGC	1923
3017	Qy	GACTCGCCACTCTCTGGTGGGATGCAACGGAAGCTGCTCGTGGGCCATTGCCCTTTGTGGG	3076
1924	Db	GAGCGCTTCTTGAGCGGGGGCATGAGGCGCAAGCTCTCCATCGGCATGCCCTCATCGC	1983
3077	Qy	CGGCTCCCAAGTTGTTATCTTGACAGAGCCTACGGCTGGCGTGGATCCTGCTTCCCGCGG	3136
1984	Db	AGGCTCCAAAGTGCTGATCTGACGACGAGCCCACTCGGGCATGSAACCCATCTCCAGGAG	2043
3137	Qy	CGGTATTGGGAGCTGCTGCTCAATACCAGAAAGTTCGACGCTGATCTCTCCACCCCA	3196
2044	Db	GGCCATCTGGGATCTTCTTTCAGCGGCAGAAAAGTGACCGCACCATCTGTGTGACCACCCA	2103
3197	Qy	CCACTGTGATGAGCAGAGTGTCTGGAGACCGGTGTGGCTGTGGTGTGGCAGAGGTGGCCGCTT	3256
2104	Db	CTTTCATGSAAGAGGCTGACCTGTCTGGGAGACCGCATCGCCATCATGTGCCCAAGGGGGAGCT	2163
3257	Qy	GTGCTGCTGGGCTCCCCACTCTTTCTTCGCGCGGTTCACCTGGGCTCCGGCTACTACCTGAC	3316
2164	Db	GCAGTGTCTGGGTCCTCGCTGTTCTCTCAAGCAGAAATACGGTGCCCGGCTATCACATGAC	2223
3317	Qy	GCTGGTGAAGG	3327
2224	Db	GCTGGTGAAGG	2234

RESULT 8

US-08-762-500-74
 ; Sequence 74, Application US/08762500
 ; Patent No. 6030806
 ; GENERAL INFORMATION:
 ; APPLICANT: Landes, Gregory M.
 ; APPLICANT: Burn, Timothy C.
 ; APPLICANT: Connors, Timothy D.
 ; APPLICANT: Dackowski, William R.
 ; APPLICANT: Van Raay, Terence J.
 ; APPLICANT: Klinger, Katherine W.
 ; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
 ; COMPOSITIONS, METHODS OF MAKING AND USING SAME
 ; NUMBER OF SEQUENCES: 83
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENZYME CORPORATION
 ; STREET: One Mountain Road
 ; CITY: Framingham
 ; STATE: Massachusetts
 ; COUNTRY: United States of America
 ; ZIP: 01701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/762,500
 ; FILING DATE: 09-DEC-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION NUMBER: US 08/665,259
 ; FILING DATE: 17-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION NUMBER: PCT/US96/10469
 ; FILING DATE: 17-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dugan, Deborah A.

```
QY 2549 -----CAGAGTCAGCCCTTGGCCCAACCCCGCTGGACCCAAAGGTGCTGGTAGAAGA 2602
Db 2075 AGGGAAGAGGAAGAAGACAGTGCACCCGAGAAAGCACTCAGAAACAGAGTACTTTGAAGC 2134
QY 2603 GGCACCGCCCGCCTGAGTCTCTGGGTATCCGTTCCGAGCCCTGGAGAGCGCTTTCCTGG 2662
Db 2135 CGAGCCAGAGACCTTGGTGGCGGGATCAAGATCAAGCACCTGTCCAAAGGTGTTCAAGGT 2194
QY 2663 AAGCCCGCA-----GCCAGCCCTCGCGGGGCTCAGCCTGGACTTCTACCAAGGCCACAT 2716
Db 2195 GGGAAATAAGACAGGCGCGCTCAGAGACTGAACCTCAACCTGTACGAGGACAGAT 2254
QY 2717 CACCGCCTTCTGGGCCCAACAGCGGGCCGCAAGACCAACCCCTGTCCATCTTTGAGTGG 2776
Db 2255 CACCGTCTGTGGGCCCAACAGCGGTGCGGGAAAGACCAACCCCTCTCCATGCTCACAGG 2314
QY 2777 CCTCTTCCACCCAGTGTGCTCTGCTTATCTTGGGCCACGACGCTCCGCTCCAGCAT 2836
Db 2315 TCTCTTTTCCCCCACCAGTGGACGGGCATACATCAGCGGTATGAAATTTTCCAGGACAT 2374
QY 2837 GCGCGCCATCGCGCCCACTGGGCGTCTGTCTCAGTACCAACGCTGTCTTTTGGATGCT 2896
Db 2375 GTTCAGATCGGAGAGCCTGGGCTGTGCCGACGACGACATCTCTGTTTGACACTT 2434
QY 2897 GACCGTGACAGACAGTCTGTGTTCTATFGGGCGGCTGAAGGCTGTGAGTGCCTGTAGT 2956
Db 2435 GACAGTCGAGACACCTTTATTTCTACGCCAGCTGAAGGGGCTGTCACTCAGAAAGTG 2494
QY 2957 GGGCCCCAGCAGACCTCTGCTGACGATGTGGGGCTGTCTTCCAGCAGAGTGTGCA 3016
Db 2495 CCCTGAAGATCAAGCAGATGCTGCACATCATCGGCCCTGGAGGACAAAGTGGAACCTCA 2554
QY 3017 GACTCGCCACCTCTCTGTGGGATGCAACGAAAGCTGTCCGTGGCCATTTGCTTTGGGG 3076
Db 2555 GAGCGCTTCTGAGCGGGGATGAGCGGACAGCTCTTCATCGCATCGCCCTCATCGC 2614
QY 3077 CGGTCTCCAAAGTTGTTATCTTGGACGACCTACCGCTGGCGTGGATCTGTTCCCGCG 3136
Db 2615 AGGTCTCAAGTGTCTACTTGGACGAGCCACCTCGGGCATGGAAGCCATCTCCAGGAG 2674
QY 3137 CGGTATTGGAGCTGTCTCAATACCGAGAGGTGCGACGCTGATCTCTCCACCCA 3196
Db 2675 GGCCATCTGGGATCTTCTTACGCGCAGAAAGTGACCGCACCATCGTGTGCTGACCA 2734
QY 3197 CCACCTGATCAGGACAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGCGCGT 3256
Db 2735 CTTATGACAGAGCTGACCTGTGGGAGACCGCATCGCCATCATGGCCAAAGGGGAGCT 2794
QY 3257 GTGCTGTGTGGCTCCCACTCTTCTGCGCGCTCACCCTGGGCTCCGCTACTACTGAC 3316
Db 2795 GCAGTGTGCGGCTCTCGCTTCTTCAAGCAGAAATACGGTGGCGGCTATCATGAC 2854
QY 3317 CTGCTGAAGG 3327
Db 2855 CTGCTGAAGG 2865
```

RESULT 9

```
US-09-724-797-7
; Sequence 7, Application US/09724797
; Patent No. 673998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 7
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(987)
US-09-724-797-7
```

Query Match 2.4%; Score 161.2; DB 4; Length 987;

```
Best Local Similarity 50.5%; Pred. No. 1.6e-24;
Matches 475; Conservative 0; Mismatches 453; Indels 12; Gaps 3;
```

```
QY 5627 GATCCAGCTGTGTGACCGCTTGTGCTGGGATTTCCCTCTGGTGTGTTTGGGCTGCT 5686
Db 60 GACCGCGCGGTGCACGCACTGGACCTGGTGGTGGGCGGGACGATCCACGGGGTCT 119
QY 5687 GGGTGTGAATGGACAGGGAAGACGTCAAGTTTCGATGTGAGCGGGGACATTTGGC 5746
Db 120 GGGCCCGAAACGGCGCCGCAAGACGACGGCCATCAAGATGCTCGCCACGCTGATGCGACC 179
QY 5747 CAGCAGGGGAGGCTGTGCTGGCAGGCAACAGAGTGGCCGGAACCCAGTGTGCGCA 5806
Db 180 CACCTCCGCAACCGCTCGTGGGCAACGCTGTCCGAGGCGCGGAGGTCCG 239
QY 5807 CCTCAGCATGGGATPACTGCTTCAATCCGATGCCATCTTTGAGTGTCTGACGGGCGCGA 5866
Db 240 GCGCGCATCGGCTCACCGGCCAGACCATGTCTCGACGAGGACATGACCGGCGTGA 299
QY 5867 GCACCTGAGCTGTGCGGCGCTGCGGGGTGTCGCGAGGCCAGGTTGCCAGACCGC 5926
Db 300 GAACCTGATCTCGCGCGCGCTGACGGGTCTGCGGCAACGCTCGCGGCGCGCGGCG 359
QY 5927 TGGCTCAGGCTGCGGCTGTGGGACTCTCATGTAGCAGACCGGCTGCGAGGACCTT 419
Db 360 GGACAGTTGATGAGGCGGTTCGACCTCACCGAGGTGCGCGCGCGCTGTTGTAAGACCTT 496
QY 5987 CAGCGGAGGAAACAAACGCAAGCTTGGCGACGGCCCTGCGCTGTGGGACCCAGCGCT 6046
Db 420 CTCGCGCGGCGCGCGCGCATCGACGTGGCGCGCGAGCATGTGTGTCACCCCGAGCT 479
QY 6047 GGTGTTTCTGGAACAGCGCGACCAAGGATGGACCCAGCGCGCGCGGCTTCTTTGAA 6106
Db 480 GCTGTTCTGACAGCGCGACCAACCGCTCTGACCCGCGAGCGCGCGAGGCTTGGGA 539
QY 6107 CAGCTTTTGGCGTGTGCGGAGGCGCTTCACTGATGCTCACCCTCCCATAGCTGGA 6166
Db 540 GATGATCCGGCGCTGTGTCGGGACCGGGGACCGTCTCTGCTGACCGAGTACCTCGA 599
QY 6167 GGAGTGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTGTGAATGGGCGGTTCCGCTGCT 6226
Db 600 CGAGCGGACCACTTCGCGCGAGCTGACCGCTCATCGACCGCGCGCATCGTGGCGCA 659
QY 6227 GGGCAGCCCGCAACATCTCAAGGGCAGATTGCGGGGGGTCAACATGACCTTCGCGGT 6286
Db 660 GGGCACCCCGCGGAGCTGAAGGGGAGCGCGCGCGCGCGG---CGTGTGCGACGTGCGGCT 716
QY 6287 GCCCGCGCAGGTTCCAGCGCGGCGGCTTCTGTCGGCGCGAGTTCCCTGGGTCGGA 6346
Db 717 GCGTGACCCCGAGCGCGCGCGGCGGCTTCTGTCGCGGAGGCGCTTCGCGCGCGCGC 776
QY 6347 GCTGCGGAGGCAATGGAGGCGCGCTTCCAGTGTGCGCGCGGAGGCGGCTGCGC 6406
Db 777 GCG---CGACCTGACCTCCGATCCGCGCGGCTGCTCGTGGGGTGACCGACCCCGACCG 833
QY 6407 CTTGGCGCGGTCTTTGAGAGCTGGCGGTGCAACGCGCGAGAGCACGCGGTGGAGGACTT 6466
Db 834 GCGGCGCTGCGCCCTGGGCGAGCTGGGCGGCGCGGCGATCCA-----CGTCGACGACTT 887
QY 6467 TTCCGTGAGCAGACGATGCTGGAGGAGGTATTCTTGTACTTCTTCAAGGACCGAGGAA 6526
Db 888 CACGCTGGCCAGCCCTCGCTCGACACGGTGTCTTCTCGCCCTCACCGGTCTACTGACGGT 947
```


Query Match	2.0%;	Score 134.2;	DB 3;	Length 4403765;
Best Local Similarity	47.8%;	Pred. No. 9.3e-18;		
Matches 456;	Conservative 0;	Mismatches 488;	Indels 9;	Gaps 2;
Qy	5570	CACCCAGGGGATGTTGGTGTGAGGAACTTGACCAAGGTATACCGTGGGACAGGAT	5629	
Db	3266492	CAACGACGACATGCGCGTGTGGTTAAACGGGGTTTCGAAGACCTTACCGCAAGGGCAAGAT	3266551	
Qy	5630	GCCAGCTGTTTACCGCTTGTGCTCTGGGGATTCCCCCTGGTGTAGTGTGTTTGGGCTGCTGGG	5689	
Db	3266552	TGTGGCCCTCGATGACGTAGTGTTCAGAGTGCGCGCGGTGAAGTATCGSGCTGCTGGG	3266611	
Qy	5690	TGTGAATGGAGCAGGGAAGAGCTCCAGTTCCTCGATGGTGAAGGGGACACATTGGCCAG	5749	
Db	3266612	CCCCAACGGGGCCGGCAAGACGACATGTGTGACATCTGTGACGCTGACCCGACCGGA	3266671	
Qy	5750	CAGGGCGAGGCTGTGCTGGCAGGCCACAGCTGCGGCCCGGGAACCCAGTCTGCGACCT	5809	
Db	3266672	TGCGGGCTCGGCGATCATCGCTGGCTAGATGTTGTTCCGAACCGGCCGGTGTACGCCG	3266731	
Qy	5810	CAGCATGGGATGATGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGCCGCGACGA	5869	
Db	3266732	CTCGATCATGCTCACCGGCGACGAGTGGCGCTCGACGACGCGCTTTCGCTGAGCAGAA	3266791	
Qy	5870	CCTGGAGCTGCTTGGCGGCTTGGCGGTGTCCGGAGGCCAGGTTGCCAGACCGCTGG	5929	
Db	3266792	CCTGGTGTGTTGGTTCGTCTGTGGGACTGAGCAAGTCCGCGCGCGCAAAACGCGCCG	3266851	
Qy	5930	CTCAGGCTTGGCGGCTCTGGGACTCTCATGTTAGCGACGCGGCTCGAGGCACCTACAG	5989	
Db	3266852	CGAAGCTCTCGAGCAATTCAGCTCTGTAATCGCGGAAGAGGCGGTGGGACCTTACTC	3266911	
Qy	5990	CGGAGGAAACAAACGCAAGCTGGCGACGGCCCTGGCGCTGTTGGGACCCAGCCGCTGGT	6049	
Db	3266912	CGGCGGAATGCGCCGACGAATAGACATCGGCTGCGGATGTTGTTGTCACACCCAGGTGGC	3266971	
Qy	6050	GTCTTCGACGAGCCGACACAGCATGGACCCGACGCGCGCGCTTCCTTTGGACAG	6109	
Db	3266972	GTCTTAGACGAGCCACACCGGCTCGATCCGAGGACGGCAAGCTATTTGGGATCT	3267031	
Qy	6110	CCTTTTGGCGTGTGCGGAGGGCCGTTTCAGTGTCTCACCTCCCATAGCATGAGGA	6169	
Db	3267032	GGTGGCCAGCTTCAAGAAGCTTGGGCATTGCCAGTTGTGACCAACGACGATCTCGAGGA	3267091	
Qy	6170	GTGTGAAGCCCTCTGCTCGGCGCTAGGCCATCATGTGTGAATGGGCGGTTCGCTGCTGGG	6229	
Db	3267092	GGCGGATGCGCTCAGTGACCGGCATCATCTGTATCGATCACGGCATATCATCCGCAAGG	3267151	
Qy	6230	CAGCCCGCAACATCTCAAGGGCAGATTGCGGGGGGTACACATGACCTGCGGGTGCC	6289	
Db	3267152	CACCGGAATGAATCAAGCACCG-----CGCGGGCACCTTCTGCGAAATAGTGCC	3267205	
Qy	6290	CGCCGCAAGTCCGACCGGACGGGCTTTCGTGGGCGCGAGTTCCTCGGGTCGAGCT	6349	
Db	3267206	CCGCGATCTGAAGGATCTGAGACGTATCTGTCGCGGCGCTCGGTTGCTGTTGCCGACGA	3267265	
Qy	6350	GCGGAGGCACATGGAGCGCTGCGCTTCCAGCTGCCCGGGAGGGCGCTGCGCCCT	6409	
Db	3267266	CCACAGGGCGATGCTGACGCGCGCATCTCAGACCCGCATTCAGATCCGCGGCTGACGGCAT	3267325	
Qy	6410	GGC---CGCGCTCTTTGGAGAGCTGGCGGTGACGGCGCAGACGACGGCTGGAGACTT	6466	
Db	3267326	ACGTATGCTCTGACGGCAGCGCCGGATTCGACGAGGCGAGGATCGAGTACCCGATAT	3267385	
Qy	6467	TTCCGTGAGCCAGACGATGTGGAGGAGGTATTCTTTGATCTTCTCCAAGACC	6519	
Db	3267386	TGGGCTCGCGCGACCGTCACTCGATACGTTATCTTGGCCATACGACACCGAT3267438		

RESULT 13

US-09-103-840A-1

Sequence 1, Application US/09103840A

```

; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

Query Match      2.0%; Score 134.2; DB 3; Length 4411529;
Best Local Similarity 47.8%; Pred. No. 9.3e-18;
Matches 456; Conservative 0; Mismatches 488; Indels 9; Gaps 2;

Qy 5570 CACCAGGGGGATGTGTGTGCTGTGAGGAACCTTGACCAAGATATACCGTGGGCGAGGAT 5629
Db 3272216 CAACGACGACATGGCGGTGTGTTAACGGGGTTTCGCAAGACCTACGCGAAGGCGAAGAT 32722755

Qy 5630 GCCAGCTGTTGACCGCTTGCTGGGATTCCCTCTGGTGAGTGTTTTGGGCTGCTGGG 5689
Db 3272276 TGTGGCCCTCGATGACGTGATGTTTCAAGGTGCGCCGCGTGAAGTATCGGGCTGCTGGG 32723355

Qy 5690 TGTGAATGAGCAGGGAAAGACGCTCCAGTTTCGCATGGTGCAGGGGACACATTTGGCCAG 5749
Db 3272336 CCCAACGGGGCGGCAAGACACCATGTTGGACATCTTGTTCGAGCTGACCCGACCGGA 32723955

Qy 5750 CAGGGCGGAGGTGTGCTGTCAGGCCACACGCTGGGCCCGGGAAACCCAGTCTCTCGCACCT 5809
Db 3272396 TGCCGGCTCGCGCATCATCGCTGGTACGATGTTGTTCCGAACCGGCCGCTGTACGCCG 32724555

Qy 5810 CAGCATGGGATCTGCGCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGCCCGGACGA 5869
Db 327456 CTGATCATGGTCAACCGGGCAGCAGGTGGCCGCTGACGACGCGCTTTCGGTGAAGCAGAA 32725155

Qy 5870 CTTGGAAGCTCTTTGCGCGCTGCGCGGTGTCGGAGGCCACAGTTGCCACAGACCCTGG 5929
Db 3273516 CTTGGTGTGTTTGGTGTGCTGTGGGGACGTGACGAGTCCGCGCGCGGCAACCGCCCGC 32725755

Qy 5930 CTCAGGCTGGCGCTCTGGACTCTCATGGTACGAGACCGCGCTTCGAGGCACTTACAG 5989
Db 3272576 CGAACTGCTCGAGCAATTCAGCCTCTGATACGCGGAAAGAGCGGTGGGCACTTACTC 32726355

Qy 5990 CGAGGGACAAACGCAAGCTGCGGACGGCCCTGGCGCTGGTTGGGGACCCAGCCGTGGT 6049
Db 3272636 CGCGGAATTCGCGCGACGAATAGACATCGCGTTCGGAATGGTGGTTCACACCCAGGTGGC 32726955

Qy 6050 GTTCTTGGACGACCGACACAGGCATGACCCACGCGCGCGCTTCTCTTTGGAACAG 6109
Db 3272696 GTTCTTAGACGACGCCACACCGGGCTCGATCCAGAGCCGCGCAAGCTATTTGGGATCT 32727555

Qy 6110 CTTTTTGGCGGTGGTGGGAGGGCGGTTTCAGTGTGCTCACCTCCCATAGCATGGAGGA 6169
Db 3272756 GGTGGCCAGCTTCAAGAGCTGGGCATTCACAGTTGTTGACCACGACGATATCTCAGGA 32728155

Qy 6170 GTGTGAAGCGCTCTGCTCGGCCCTAGCGATCATGTGTGAATGGCGGGTTCCGCTGCTGGG 6229
Db 3272816 GGGCGGATGGCGTCAGTGAACCGCATCATCTCTGATCGATTCAGGCAATATCATCGCCGAAG 32728755

Qy 6230 CAGCCCGCAACATCTCAAGGGCAGATTTCGGCGGGGTACACACTGACCTCGCGGTGCC 6289
Db 3272876 CACCGGAATGAATCAAGACCG-----CGCCGCGACACCTTCTGCGAATAGTGCC 32729295

Qy 6290 CGCCGCAAGTCCACGCCGCGCAGCGGCTTTCGTGGCGGCGCGAGTTCCCTGGGTTCGAGCT 6349

```

Db 3272930 CCGGATCTGAAGATCTGACGCTATCTGTCGGCGCTCGGTTGCTGTTGCCGAGCA 3272989
QY 6350 GCGGAGGACATGAGAGCGCGCTCGCTTCAGCTGCGCGCCGAGAGGGCGCTGCGCCCT 6409
Db 3272990 CCACAGGGCGATCTGACGCCGACTCAGACCGCATTTACGATCCGCGCGCTGACGGCAT 3273049
QY 6410 GGC---GCGGCTTTTGGAGAGCTGGCGTGACGGGACAGACACGGCGTGGAGACATT 6466
Db 3273050 ACCTATGCTCGTACGAGCAGCGCGCGGATCGACGAGCGAGGATCGAGTACCGCATAT 3273109
QY 6467 TTCGCTGAGCCAGCATGCTGAGGAGGTATTCTTGTACTTCTTCCAAAGGACC 6519
Db 3273110 TCGCTGCGCGCGACCGCTCACTCGATCAGGTATTCTTGGCCATGACACCGATC 3273162

RESULT 14
US-09-252-991A-4588/c
; Sequence 4588, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4588
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4588

Query Match 1.8%; Score 125.2; DB 4; Length 999;
Best Local Similarity 48.5%; Pred. No. 5.5e-17;
Matches 377; Conservative 0; Mismatches 398; Indels 3; Gaps 1;

QY 2676 GCCTTGGGGGGCTCAGCCTGGACTTCTACAGGGCCACATCACCGCTTCTCTGGGCCAC 2735
Db 939 GCCTCAAGGGCATCGACCTGGACGTCGCGAAGGTGATTTCTTCCGCTTGTCTGGCCCC 880
QY 2736 AACGGGGCGGCAAGACCAACCCCTGTCCATCTTGAGTGGCTCTTCCACCCAGTGT 2795
Db 879 AACGGCGGGCAAGTCCACCACTCGGATTTCTTCGACCTGTGTGAACAAGACGAGC 820
QY 2796 GGCTTGCCTTATCTTGGGCGAGGAGTCCGCTCCAGATGCGCGGCATCCGGCCCCAC 2855
Db 819 GGTTCGCTCTCGTGTTCGGCGCCAGACCTCGAAGACCCGCGCGCTCAAGCGTTGC 760
QY 2856 CTGGGCTGTCTCTAGTACACGCTGTCTTTGATCATGTGACCGTGGACGAGCAGTC 2915
Db 759 CTTGGGCTGTGCGCAGGAGTTCACTTCAACAGTTTCGAGAGGTCTTCGACATGTC 700
QY 2916 TGGTTCTATGGGCGGTGAAGGCTCTAGTCCGCTGTAGTGGGCCCCGAGAGGACCGT 2975
Db 699 GTGACCCAGCGGCTACTACGGCATCCGCGGCAAGATCGCAAGGAGCGCGCGAGGC 640
QY 2976 CTGCTGAGGATGTGGGCTGTGTTCCAGCAGAGTGTGAGACTCGCACCTCTCTGTT 3035
Db 639 TACCTGACCCAACTGGGCGCTGTGGGCAAGCGAACGAGCCCTCGCGGATGCTCTCCGGC 580
QY 3036 GGGATGCAACGGAAGCTGTCCGTGGCATTTGCTTTGTGGCGGCTCCCAAGTTGTTATC 3095
Db 579 GGCATGAGCGGCGCTGATGATCGCCGCGCATGCTGTCACAGCGCGCTCTGCTGATC 520
QY 3096 CTGGACGAGCTTACGGCTGCGGTGATTCCTGTTCCCGCGCGGTATTTGGGAGCTGCTG 3155
Db 519 CTGACGAGGCCACCGCGGGGTGACATCGAGCTGCTGTTTCGATGTGGAGCTTCTCTC 460

QY 3156 CTCAAA---TACGAGAAGGTCCGACGCTGATCTCTCCACCACCACCTGGATGAGGCA 3212
Db 459 ACAGAACTCAACACAGGAAGGCATCAGCATATCTCTCACCACCACCTACTCTGAGAGGGC 400
QY 3213 GAGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGTGGCTCC 3272
Db 399 GAGCAGCTCTGCGGCAACATTGCCATCATCGACACCGGCGAGATCGTGAGAACACGAGC 340
QY 3273 CCACCTTCTTCTGGGCGGTCACTGGCTCGGCTACTACTGACGCTGTGTGAGGGCCGC 3332
Db 339 ATGCGCGACTCTGTGATGACCTCTGATTCGGAACCTTCTCTGCTGACCTGAAGAAGCTC 280
QY 3333 CTGCCCCCTCACCAACAATGAGAAGGTGACATGACATGAGGGGAGTGTGACACGAGG 3392
Db 279 CAGGCCCTGCCGCGACCTCGACGGCTATCGACGGGCTGGTGACACACACTCTC 220
QY 3393 CAGGAAAGAAGATGGCAGCCAGGACGAGTGGCACTCTCTCAGCTGTGGCCCC 3450
Db 219 GAAGTGCAGTGGAGAAGAGCCAGGCGCATCAACGACCTGTTCGCCCACTGGGTGCAC 162

RESULT 15
US-09-252-991A-4266
; Sequence 4266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4266
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4266

Query Match 1.8%; Score 125.2; DB 4; Length 1008;
Best Local Similarity 48.5%; Pred. No. 5.6e-17;
Matches 377; Conservative 0; Mismatches 398; Indels 3; Gaps 1;

QY 2676 GCCTTGGGGGGCTCAGCCTGGACTTCTACAGGGCCACATCACCGCTTCTCTGGGCCAC 2735
Db 133 GCCTCAAGGGCATCGACCTGGACGTCGCGAAGGTGATTTCTTCCGCTTGTCTGGCCCC 192
QY 2736 AACGGGGCGGCAAGACCAACCCCTGTCCATCTTGAGTGGCTCTTCCACCCAGTGT 2795
Db 193 AACGGCGGGCAAGTCCACCACTCGGATTTCTTCGACCTGTGTGAACAAGACGAGC 252
QY 2796 GGCTTGCCTTATCTTGGGCGAGGAGTCCGCTCCAGATGCGCGGCATCCGGCCCCAC 2855
Db 253 GGTTCGCTCTCGTGTTCGGCGCCAGACCTCGAAGGACCCGCGCGCTCAAGCGTTGC 312
QY 2856 CTGGGCGTCTGTCTCAGTACAAACGCTGTGTTGATCATGTGACCGTGGACGAGCAGTC 2915
Db 313 CTTGGGCTGTGCGCAGGAGTTCACTTCAACAGTTTCGAGAGGTCTTCGACATGCTC 372
QY 2916 TGGTTCTATGGGCGGTGAAGGCTCTGAGTGGCGCTGTAGTGGGCCCCGAGAGGACCGT 2975
Db 373 GTGACCCAGCGGCTACTACGGCATCCGCGGCAAGATCGCAAGGAGCGCGCGAGCGC 432
QY 2976 CTGCTGAGGATGTGGGCTGTCTCCAAAGCAGAGTGTGAGACTCGCACCTCTCTGTT 3035
Db 433 TACCTGACCCAACTGGGCGCTGTGGGCAAGCGCAACGAGCCCTCGCGGATGCTCTCGGC 492
QY 3036 GGGATGCAACGGAAGCTGTCCGTTGGGCGCATTTGCTTTGTGGGCGGCTCCCAAGTTGTTATC 3095

[illegible]

Search completed: December 29, 2004, 23:55:32
Job time : 575 secs

This Page Blank (Uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 23:47:03 ; Search time 3209 Seconds
(without alignments)
11975.922 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 CTCAGGGCGGCGCGTCCCC.....CGTGTGGTGAAGAAAAA 6804

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	6804	100.0	6804	9	US-09-995-542-4
2	6762.4	99.4	6791	16	US-10-332-447-59
3	6717.8	98.7	6768	9	US-09-858-194-1
4	6717.8	98.7	6768	15	US-10-154-419-1
5	6710	98.6	6795	10	US-09-983-446A-8
6	6686.4	98.3	6704	17	US-10-775-920-1
7	6684.8	98.2	6704	17	US-10-775-920-3
8	6565.6	96.5	6588	17	US-10-775-920-4
9	6478.4	95.2	6522	16	US-10-182-006-1
10	6398.4	94.0	6432	9	US-09-858-194-3
11	6398.4	94.0	6432	15	US-10-154-419-3
12	6073.6	89.3	6174	17	US-10-775-920-6

13	6036.8	88.7	6327	16	US-10-114-270-175
14	5928.2	87.1	6027	17	US-10-775-920-5
15	5458.2	80.2	5669	16	US-10-182-006-3
16	4381.6	64.4	4653	9	US-09-995-542-7
17	3902.4	57.4	4633	9	US-09-995-542-1
18	3447.4	50.7	3635	9	US-09-767-870-3
19	3447.4	50.7	3635	15	US-10-242-568-3
20	3081.8	45.3	3609	9	US-09-833-381-958
21	1899.2	27.9	2075	9	US-09-833-381-962
22	1808.6	26.6	1879	16	US-10-108-260A-1831
23	1808.6	26.6	1879	17	US-10-775-920-7
24	1740.6	25.6	6786	15	US-10-313-641-8
25	1740.6	25.6	6786	15	US-10-428-551-8
26	1739	25.6	10412	15	US-10-313-641-7
27	1739	25.6	10412	15	US-10-428-551-7
28	1737.4	25.5	7860	16	US-10-452-510-2
29	1737.4	25.5	7860	16	US-10-617-334-2
30	1737.4	25.5	7860	17	US-10-745-377-6
31	1737.4	25.5	7860	17	US-10-444-465-2
32	1737.4	25.5	7860	17	US-10-833-679-2
33	1737.4	25.5	7860	18	US-10-872-113-6
34	1737.4	25.5	10365	14	US-10-098-939-1
35	1737.4	25.5	10381	14	US-10-098-939-2
36	1737.4	25.5	10423	14	US-10-098-939-3
37	1735.8	25.5	9741	10	US-09-984-827-1
38	1735.8	25.5	9870	10	US-09-984-827-106
39	1735.8	25.5	9870	10	US-09-984-827-107
40	1735.8	25.5	9870	10	US-09-984-827-109
41	1735.8	25.5	9870	10	US-09-984-827-110
42	1734.2	25.5	9497	15	US-10-170-385-294
43	1734.2	25.5	9497	16	US-10-429-160-3
44	1734.2	25.5	9497	17	US-10-465-498-91
45	1734.2	25.5	9741	9	US-09-846-456-10

ALIGNMENTS

RESULT 1
US-09-995-542-4
; Sequence 4, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (210)..(347)
; NAME/KEY: CDS
; LOCATION: (210)..(6650)
US-09-995-542-4

Query Match 100.0%; Score 6804; DB 9; Length 6804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	1	CTCAGGGCGGCGCGTCCCTGCTGCTGCTGCGGAGGCGGCGGAGAGCTGCGG	60
Db	1	CTCAGGGCGGCGCGCTCCCTGCTGCTGCTGCGGAGGCGGAGGCGGAGAGCTGCGG	60

Qy	61	AGCCCTCGGAAGAGCTTCCAGGAACCCCTGCGCTGTGGGATAAAGGAATAGGTTCAAAA	120
Db	61	AGCCCTCGGAAGAGCTTCCAGGAACCCCTGCGCTGTGGGATAAAGGAATAGGTTCAAAA	120
Qy	121	GGGCGAGGGAGTTGCCGCGACCGCACCGTCTTCAGCCGACCGTGTCTGACCT	180
Db	121	GGGCGAGGGAGTTGCCGCGACCGCACCGTCTTCAGCCGACCGTGTCTGACCT	180
Qy	181	CTCTGTCCCGTCCCGCCAGTCTCACCATGGCTTCTGGACACAGCTGATGCTGCTGC	240
Db	181	CTCTGTCCCGTCCCGCCAGTCTCACCATGGCTTCTGGACACAGCTGATGCTGCTGC	240
Qy	241	TCTGGAAAGAAATTTTCATGTATCGCGGAGACAGCGGTCACAGCTCTCTGGTCEAAATGCTGT	300
Db	241	TCTGGAAAGAAATTTTCATGTATCGCGGAGACAGCGGTCACAGCTCTCTGGTCEAAATGCTGT	300
Qy	301	GGCCTCTCTTCTCTTCTTATCCTGTGTGCTGTTCGCGACTCTCCACCCGCCCTGTGGAGC	360
Db	301	GGCCTCTCTTCTCTTCTTATCCTGTGTGCTGTTCGCGACTCTCCACCCGCCCTGTGGAGC	360
Qy	361	ACCATGAATGCACTTCCCAAAACAAGCACCTGCCATCGCGGCGACCGTGCCTGGCTCC	420
Db	361	ACCATGAATGCACTTCCCAAAACAAGCACCTGCCATCGCGGCGACCGTGCCTGGCTCC	420
Qy	421	AGGTCCTCATCTGTAAATGTGAACAACACCTGCTTTCCGACAGCTGACACCGGGCGAGGAGC	480
Db	421	AGGTCCTCATCTGTAAATGTGAACAACACCTGCTTTCCGACAGCTGACACCGGGCGAGGAGC	480
Qy	481	CGGGCGCCTGAGCAACTTCAAAGACTCCTTGTCTCCCGCTGTAGCCGATGCCGCA	540
Db	481	CGGGCGCCTGAGCAACTTCAAAGACTCCTTGTCTCCCGCTGTAGCCGATGCCGCA	540
Qy	541	CTGTGCTGGAGGGGCGAGTGCCACAGGACGCTGGCTGGCTTAGGGAAGCTGATCGCCA	600
Db	541	CTGTGCTGGAGGGGCGAGTGCCACAGGACGCTGGCTGGCTTAGGGAAGCTGATCGCCA	600
Qy	601	CGCTGAGGGCTGACGACGACGCGCCAGGCTCAACCAACAAAGCAGTCTCCACTGGAAAC	660
Db	601	CGCTGAGGGCTGACGACGACGCGCCAGGCTCAACCAACAAAGCAGTCTCCACTGGAAAC	660
Qy	661	CACCCATGCTGGATGTGCGGAGCTGTGACGTCACTGTCTGCGGACGGAAATCCCTGGGGT	720
Db	661	CACCCATGCTGGATGTGCGGAGCTGTGACGTCACTGTCTGCGGACGGAAATCCCTGGGGT	720
Qy	721	TGGCACCTGGGCGAAGCCAGGAGCCCTTGACAGCTTTGTTGGAGGCGCTGAGGACCTGG	780
Db	721	TGGCACCTGGGCGAAGCCAGGAGCCCTTGACAGCTTTGTTGGAGGCGCTGAGGACCTGG	780
Qy	781	CCCAGAGCTCTCTGGCGCTGCGACGCTGTGGAGCTTCGGGCACTGCTGCAGAGACCCC	840
Db	781	CCCAGAGCTCTCTGGCGCTGCGACGCTGTGGAGCTTCGGGCACTGCTGCAGAGACCCC	840
Qy	841	GAGGGAACAGCGGCCCTGTGAGTTGTCTCAGAGGCCCTCTGCAGTGTGAGGGACCTTA	900
Db	841	GAGGGAACAGCGGCCCTGTGAGTTGTCTCAGAGGCCCTCTGCAGTGTGAGGGACCTTA	900
Qy	901	GCAGCAAGTGGGCCCTCTCCTCACTGTTAGAGGCTAGTGAACCTGAGCTGGTGG	960
Db	901	GCAGCAAGTGGGCCCTCTCCTCACTGTTAGAGGCTAGTGAACCTGAGCTGGTGG	960
Qy	961	GGCAGGAGCCAGAAATCGCGCCCTGCACAGCAGCTTGAGCCCGCTGTCTGGAGCTGA	1020
Db	961	GGCAGGAGCCAGAAATCGCGCCCTGCACAGCAGCTTGAGCCCGCTGTCTGGAGCTGA	1020
Qy	1021	TTGAGGCCCTTGGAACAGCAACCCGTGTCCCGCTGTCTTGAGACGCTCTGAACGCTCTGA	1080
Db	1021	TTGAGGCCCTTGGAACAGCAACCCGTGTCCCGCTGTCTTGAGACGCTCTGAACGCTCTGA	1080
Qy	1081	TCCTCGGGAAGCTACTCTTTTGCAACAGATACACTTTTACCCGGAAGCTCATGGCCAGG	1140
Db	1081	TCCTCGGGAAGCTACTCTTTTGCAACAGATACACTTTTACCCGGAAGCTCATGGCCAGG	1140
Qy	1141	TGAACCCGACCTTCGAGGAGCTACCCCTGTGTGAGGATGTCCGGGAGGCTGTGGGAGATGC	1200

Db	1141	 TGAACCGGACCTTCGAGGAGCTCACCTCTGCTGAGGATGTCCGGGAGGTGTGGGATGC	1200
Qy	1201	TGGACCCCGGATCTTCACCTTCATGAACACAGTGTCCAAATGTGGCCATGCTCTCAGCGGC	1260
Db	1201	TGGGACCCCGGATCTTCACCTTCATGAACACAGTGTCCAAATGTGGCCATGCTCTCAGCGGC	1260
Qy	1261	TCCTGCAGATGCAAGGATGAAGGAAGGCAGCCACAGACCTGGAGGCCGGACCATGG	1320
Db	1261	TCCTGCAGATGCAAGGATGAAGGAAGGCAGCCACAGACCTGGAGGCCGGACCATGG	1320
Qy	1321	AGGCCCTGCATCCTTTCTGGACCCCTGGGAGCGGTGGCTACAGCTGGCAGGACGACACG	1380
Db	1321	AGGCCCTGCATCCTTTCTGGACCCCTGGGAGCGGTGGCTACAGCTGGCAGGACGACACG	1380
Qy	1381	CTGATGTGGGGCACCTGTGTGGGACCGCTGGGGCGAGTGACGGAGTGCTGTCTCTGGACA	1440
Db	1381	CTGATGTGGGGCACCTGTGTGGGACCGCTGGGGCGAGTGACGGAGTGCTGTCTCTGGACA	1440
Qy	1441	AGCTGGAGGCGGACCCCTCAGAGGACGCCCTGTGTGCGGGGCCCTGCAACTGCTCGCGG	1500
Db	1441	AGCTGGAGGCGGACCCCTCAGAGGACGCCCTGTGTGCGGGGCCCTGCAACTGCTCGCGG	1500
Qy	1501	AA CATCGATTCTGGGCGGGGTGCTTTCTTTGGGAACTTGAGGACTCTTTAGACCCACAG	1560
Db	1501	AA CATCGATTCTGGGCGGGGTGCTTTCTTTGGGAACTTGAGGACTCTTTAGACCCACAG	1560
Qy	1561	AGCACCCAAACCCAGACACTGGGGCCCGGCCACGTGGGCATCAAAATCCGCATGACATTG	1620
Db	1561	AGCACCCAAACCCAGACACTGGGGCCCGGCCACGTGGGCATCAAAATCCGCATGACATTG	1620
Qy	1621	ACGTGTGTCAGGAGACCAATAAGATCAGGAGACAGTGTGGGACCTTGGGCCACAGCCGCG	1680
Db	1621	ACGTGTGTCAGGAGACCAATAAGATCAGGAGACAGTGTGGGACCTTGGGCCACAGCCGCG	1680
Qy	1681	ACCCCTTGACCGACTCTGCGCTACGTGTGGGGCGGCTTCTGTGTAACCTGCAAGACCTGGTGG	1740
Db	1681	ACCCCTTGACCGACTCTGCGCTACGTGTGGGGCGGCTTCTGTGTAACCTGCAAGACCTGGTGG	1740
Qy	1741	AGCGTCAGCGCTCCGCTGCTCAGCGGCGCCAAACCCCGGGCCGGCTCTACCTGCAGC	1800
Db	1741	AGCGTCAGCGCTCCGCTGCTCAGCGGCGCCAAACCCCGGGCCGGCTCTACCTGCAGC	1800
Qy	1801	AGATGCCCTATCCGTGCTATGTGGAGGACGTGTTCTCGTGTGCTGAGCCGCTGCTGC	1860
Db	1801	AGATGCCCTATCCGTGCTATGTGGAGGACGTGTTCTCGTGTGCTGAGCCGCTGCTGC	1860
Qy	1861	CGCTCTCTCTGACGCTGGCTGATCTA CTCCGTGACACTGACAGTGAAGGCCGCTGCTGC	1920
Db	1861	CGCTCTCTCTGACGCTGGCTGATCTA CTCCGTGACACTGACAGTGAAGGCCGCTGCTGC	1920
Qy	1921	GGGAGGAGGACGCGGCTCGGGACCAACATGCGGCCATGGGGCTCAGCCGCGCGGTGC	1980
Db	1921	GGGAGGAGGACGCGGCTCGGGACCAACATGCGGCCATGGGGCTCAGCCGCGCGGTGC	1980
Qy	1981	TCCTGGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCTGTCTCAGCCGCGCTGCTGG	2040
Db	1981	TCCTGGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCTGTCTCAGCCGCGCTGCTGG	2040
Qy	2041	TTCTGTGTCTCAAGCTGGGGACATCTCTCCCTACAGCCACCCGGGGTGGTCTTCTGT	2100
Db	2041	TTCTGTGTCTCAAGCTGGGGACATCTCTCCCTACAGCCACCCGGGGTGGTCTTCTGT	2100
Qy	2101	TCCTTGGCAGCTTCGGGTGGCCACGGTGACACAGAGCTTCCTGTCTCAGCGCCTCTTCT	2160
Db	2101	TCCTTGGCAGCTTCGGGTGGCCACGGTGACACAGAGCTTCCTGTCTCAGCGCCTCTTCT	2160
Qy	2161	CCGCGCCAACTGGGTGCGGCTCGGGCGGCTGCGCTACTTCTCCCTACTACTGCGCT	2220
Db	2161	CCGCGCCAACTGGGTGCGGCTCGGGCGGCTGCGCTACTTCTCCCTACTACTGCGCT	2220
Qy	2221	ACGTGTGTGTGTGGCTTGGCGGGAACGGGTGCGCGGGGTGGCGGCTGGCGGAGCC	2280

Db 2221 ACCTGCTGTGTGGCTTTGGCGGACCGGCTGCCCGCGGTGGCCGCGTGGCGCGGAGCC 2280
Qy 2281 TGCTGTCCCGTGGCTTTGGCTTTGGCTTGGAGAGCTGGCTGTCTGTGGAGAGCAGG 2340
Db 2281 TGCTGTCCCGTGGCTTTGGCTTTGGCTTGGAGAGCTGGCTGTCTGTGGAGAGCAGG 2340
Qy 2341 GCGAGGCGCGAGTGGGACAAAGTGGGACCCGGCCTACGGCAGACGTCTTACGCTGG 2400
Db 2341 GCGAGGCGCGAGTGGGACAAAGTGGGACCCGGCCTACGGCAGACGTCTTACGCTGG 2400
Qy 2401 CCCAGTCTCTGGCCCTTCTGTCTGTGGACGGGCGCTCTACGGCCTGCCACCTGGTACC 2460
Db 2401 CCCAGTCTCTGGCCCTTCTGTCTGTGGACGGGCGCTCTACGGCCTGCCACCTGGTACC 2460
Qy 2461 TGGAGCTGTGTGCCCAGGCCAGTACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Db 2461 TGGAGCTGTGTGCCCAGGCCAGTACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Qy 2521 GGAGCTACTGTGGGACCTTGGCCCGCCCAAGAGTCCAGCCCTTGGCCCAACCCGCTGG 2580
Db 2521 GGAGCTACTGTGGGACCTTGGCCCGCCCAAGAGTCCAGCCCTTGGCCCAACCCGCTGG 2580
Qy 2581 ACCCAAAGTCTGTGTAGAGGACCGCCGCGCTGAGTCTGGCGTATCCGTTGCA 2640
Db 2581 ACCCAAAGTCTGTGTAGAGGACCGCCGCGCTGAGTCTGGCGTATCCGTTGCA 2640
Qy 2641 GCCTGGAGAAGCGCTTTCTTGGAGCCCGCAGCAGCCCTCGGGGGCTCAGCCTGGACT 2700
Db 2641 GCCTGGAGAAGCGCTTTCTTGGAGCCCGCAGCAGCCCTCGGGGGCTCAGCCTGGACT 2700
Qy 2701 TCTACAGGGCCACATACCGCTTCTTGGGCCAACAGCGGCGCGGCAAGCACCAACC 2760
Db 2701 TCTACAGGGCCACATACCGCTTCTTGGGCCAACAGCGGCGCGGCAAGCACCAACC 2760
Qy 2761 TGTCCATCTTCAAGTGGCTTCTTCCACCAAGTGGTGGCTTCACTTCTTGGGCCACG 2820
Db 2761 TGTCCATCTTCAAGTGGCTTCTTCCACCAAGTGGTGGCTTCACTTCTTGGGCCACG 2820
Qy 2821 ACGTCCGCTCAGATGAGCGCCATCCCGGCCACCTGGGCGTGTCTCCTCAGTACAAAG 2880
Db 2821 ACGTCCGCTCAGATGAGCGCGCATCCCGGCCACCTGGGCGTGTCTCCTCAGTACAAAG 2880
Qy 2881 TGCTGTTTGAATGCTGACCGTGGACGAGCAGTCTGGTTCTATGGGCGGCTGAAGGTC 2940
Db 2881 TGCTGTTTGAATGCTGACCGTGGACGAGCAGTCTGGTTCTATGGGCGGCTGAAGGTC 2940
Qy 2941 TGAGTGGCGCTGTAGTGGGCGCCGAGCAGGACCGTCTGCTCAGGATGTGGGCTGTCT 3000
Db 2941 TGAGTGGCGCTGTAGTGGGCGCCGAGCAGGACCGTCTGCTCAGGATGTGGGCTGTCT 3000
Qy 3001 CCAAGCAGTGTGAGACTCGCACTCGCACTCTCTGTGGGATGCAACGGAGCTGTCCGTGG 3060
Db 3001 CCAAGCAGTGTGAGACTCGCACTCTCTGTGGGATGCAACGGAGCTGTCCGTGG 3060
Qy 3061 CCAATTGCTTTTGGGCGGCTCCCAAGTTTATCTTGGAGAGCTACGGCTGGCGTGG 3120
Db 3061 CCAATTGCTTTTGGGCGGCTCCCAAGTTTATCTTGGAGAGCTACGGCTGGCGTGG 3120
Qy 3121 ATCTGCTTCCCGCGCGTATTTGGGAGCTGTCTGCTCAATACCGAAGAGTGGCAGCG 3180
Db 3121 ATCTGCTTCCCGCGCGTATTTGGGAGCTGTCTGCTCAATACCGAAGAGTGGCAGCG 3180
Qy 3181 TGATCTCTCCACCCACCTGATGAGGAGAGCTGTGGGAGACCGTGTGGCTGTGG 3240
Db 3181 TGATCTCTCTCCACCCACCTGATGAGGAGAGCTGTGGGAGACCGTGTGGCTGTGG 3240
Qy 3241 TGGCAGTGGCGCTTGTGCTGTGGCTTCCCACTTCTTCTGCGCGCTCACCTGGGCT 3300
Db 3241 TGGCAGTGGCGCTTGTGCTGTGGCTTCCCACTTCTTCTGCGCGCTCACCTGGGCT 3300
Qy 3301 CCGGCTACTACTGACGCTGTGTAAGGCCCGCTGCCCCCTGACCAACCAATGAGAGGCTG 3360
Db 3301 CCGGCTACTACTGACGCTGTGTAAGGCCCGCTGCCCCCTGACCAACCAATGAGAGGCTG 3360

Qy 3361 AACTGACATGAGGCGCAGTGTGGACACAGGCAAGAAAGAAATGGCAGCCAGGCA 3420
Db 3361 AACTGACATGAGGCGCAGTGTGGACACAGGCAAGAAAGAAATGGCAGCCAGGCA 3420
Qy 3421 GCAGAGTGGCAGCTCTCTCAGTGTGGCCCTGGTACAGCACTGGGTGCCCGGGCACGGC 3480
Db 3421 GCAGAGTGGCAGCTCTCTCAGTGTGGCCCTGGTACAGCACTGGGTGCCCGGGCACGGC 3480
Qy 3481 TGGTGGAGGAGCTGCCACACAGAGTGGTGTGGTGTGCCCTACACGGGTGCCCATGACG 3540
Db 3481 TGGTGGAGGAGCTGCCACACAGAGTGGTGTGGTGTGCCCTACACGGGTGCCCATGACG 3540
Qy 3541 GCAGCTTTCGCCCACTCTTCCGAGAGCTAGACACGGCTGGCGAGCTGAGGCTCACTG 3600
Db 3541 GCAGCTTTCGCCCACTCTTCCGAGAGCTAGACACGGCTGGCGAGCTGAGGCTCACTG 3600
Qy 3601 GCTACGGGATCTCCGACACAGAGTGGGATGCTTCTGAAAGTGGTGGAGAGTGTG 3660
Db 3601 GCTACGGGATCTCCGACACAGAGTGGGATGCTTCTGAAAGTGGTGGAGAGTGTG 3660
Qy 3661 CTGCGGACACAGATATGAGGATGCGAGCTGCGGCGCAGCACTATGACAGGCAATTGCTG 3720
Db 3661 CTGCGGACACAGATATGAGGATGCGAGCTGCGGCGCAGCACTATGACAGGCAATTGCTG 3720
Qy 3721 GCCTAGACGTAAACCTTCGGCTCAAGATGCCGCCACAGGAGACAGCGCTGGAGAACGGG 3780
Db 3721 GCCTAGACGTAAACCTTCGGCTCAAGATGCCGCCACAGGAGACAGCGCTGGAGAACGGG 3780
Qy 3781 AACAGCTGGGTAGCCCCAGAGACTGACAGGCTCTGGGCCAGACCGCTGGGCCGGG 3840
Db 3781 AACAGCTGGGTAGCCCCAGAGACTGACAGGCTCTGGGCCAGACCGCTGGGCCGGG 3840
Qy 3841 TACAGGCTGGGCACTGACCCGCGCAGAGCTTCCAGGCGCTCTTCTCAAGCGCTTTCTGC 3900
Db 3841 TACAGGCTGGGCACTGACCCGCGCAGAGCTTCCAGGCGCTCTTCTCAAGCGCTTTCTGC 3900
Qy 3901 TTGCCCGCGCAGCGCGCGCGCTGTTCGCCCAGATCGTCTGCTGCCCTCTTTGTGG 3960
Db 3901 TTGCCCGCGCAGCGCGCGCGCTGTTCGCCCAGATCGTCTGCTGCCCTCTTTGTGG 3960
Qy 3961 GCCTGGCCCTCTGTTCAGGCTCATGCTGCTCTTTTGGGCACTACCCGGCTCTGGCGC 4020
Db 3961 GCCTGGCCCTCTGTTCAGGCTCATGCTGCTCTCTTTTGGGCACTACCCGGCTCTGGCGC 4020
Qy 4021 TCAGTCCCAACATGACGGTGTCTCAGGTGTCTTCTTCAAGTGGAGCGCCCGAGGAGCC 4080
Db 4021 TCAGTCCCAACATGACGGTGTCTCAGGTGTCTTCTTCAAGTGGAGCGCCCGAGGAGCC 4080
Qy 4081 CTGAGCGTGGCCGCGCTCTCGAGCGCTGTTCAGGAGGAGGAGCTGGAGAGCGCCCGAG 4140
Db 4081 CTGAGCGTGGCCGCGCTCTCGAGCGCTGTTCAGGAGGAGGAGCTGGAGAGCGCCCGAG 4140
Qy 4141 TGCAGCATAGCTCCACAGGTTCTCGGACACAGAGTTCTCTGCTGAAAGTGGCCAAAGTCT 4200
Db 4141 TGCAGCATAGCTCCACAGGTTCTCGGACACAGAGTTCTCTGCTGAAAGTGGCCAAAGTCT 4200
Qy 4201 TGGCCAGTGGCACTGACCCAGAGTCTCCATCCCGAGCGCTGCTGCTGAGGAGCGCCG 4260
Db 4201 TGGCCAGTGGCACTGACCCAGAGTCTCCATCCCGAGCGCTGCTGCTGAGGAGCGCCG 4260
Qy 4261 GTGCCCGGCGCTGCTGCCCGAGCTGCGGCTGAGCTGGTGTGCCCTCCCGCCCGAGG 4320
Db 4261 GTGCCCGGCGCTGCTGCCCGAGCTGCGGCTGAGCTGGTGTGCCCTCCCGCCCGAGG 4320
Qy 4321 CAGTACCGGCTCTGGGGAAGTGGTTCAGAACCTGACAGGCGGAACTGCTGCTGACTTCC 4380
Db 4321 CAGTACCGGCTCTGGGGAAGTGGTTCAGAACCTGACAGGCGGAACTGCTGCTGACTTCC 4380
Qy 4381 TGGTCAAGACCTTACCCCGGCTGTGGCGCAGGCGCTGAGAGCTAAGAGTGGGTGAATG 4440
Db 4381 TGGTCAAGACCTTACCCCGGCTGTGGCGCAGGCGCTGAGAGCTAAGAGTGGGTGAATG 4440

QY	4441	AGGTACAGTACGAGGCTTCTCGTGGGGGCGAGACACGAGCCCTGCGCTCGGGCAAG	4500
Db	4441	AGGTACAGTACGAGGCTTCTCGTGGGGGCGAGACACGAGCCCTGCGCTCGGGCAAG	4500
QY	4501	AGTTGGGCGGCTCAGTGGAGAGTTGTGGCGCTGCTGAGTCCCTGCTGGCGGGCCC	4560
Db	4501	AGTTGGGCGGCTCAGTGGAGAGTTGTGGCGCTGCTGAGTCCCTGCTGGCGGGCCC	4560
QY	4561	TCGACCGTGTCTGTAAGAACTCAGAGCTGGGCTCAGAGCTGGATGCTCAGAGCAGTC	4620
Db	4561	TCGACCGTGTCTGTAAGAACTCAGAGCTGGGCTCAGAGCTGGATGCTCAGAGCAGTC	4620
QY	4621	TCAGATCTGGTTCACAAACAAAGGCTGGCACTCCATGTGGGCTTTGTCAACCGAGCCA	4680
Db	4621	TCAGATCTGGTTCACAAACAAAGGCTGGCACTCCATGTGGGCTTTGTCAACCGAGCCA	4680
QY	4681	GCAACGCAATCTCCGCTGCTCACTGCCCGCAGGCGCGGCCCGCCACAGCATCA	4740
Db	4681	GCAACGCAATCTCCGCTGCTCACTGCCCGCAGGCGCGGCCCGCCACAGCATCA	4740
QY	4741	CCACACTCAACCAACCCCTTGAACTCCTCAAGAGAGCAGTGTCTGAGGCTGCACTGATGG	4800
Db	4741	CCACACTCAACCAACCCCTTGAACTCCTCAAGAGAGCAGTGTCTGAGGCTGCACTGATGG	4800
QY	4801	CCTCCTCGGTGGAGCTCCTCGTCTCCATCTGTGTGTTTGGCATGTCTTGTGCCGG	4860
Db	4801	CCTCCTCGGTGGAGCTCCTCGTCTCCATCTGTGTGTTTGGCATGTCTTGTGCCGG	4860
QY	4861	CCAGCTTCACTCTGTCTCATTTGAGGAGGAGTCAACCCAGGCAAGCAGCTCAGCTCA	4920
Db	4861	CCAGCTTCACTCTGTCTCATTTGAGGAGGAGTCAACCCAGGCAAGCAGCTCAGCTCA	4920
QY	4921	TGGGGGCGCTGTCCCGACCTCTACTGCTGGGCACTTTCTCTGGGCACTGTGTAAC	4980
Db	4921	TGGGGGCGCTGTCCCGACCTCTACTGCTGGGCACTTTCTCTGGGCACTGTGTAAC	4980
QY	4981	ACTTGTGCCAGATGATCGTGTGCTCATCTTTCTGGGCTTCCAGCAGAGGCAATG	5040
Db	4981	ACTTGTGCCAGATGATCGTGTGCTCATCTTTCTGGGCTTCCAGCAGAGGCAATG	5040
QY	5041	TGGCCCTGCAACCTGCTGCTCTCTGCTGTGTGCTACTGATGCTGCTGCTGATCA	5100
Db	5041	TGGCCCTGCAACCTGCTGCTCTCTGCTGTGTGCTACTGATGCTGCTGCTGATCA	5100
QY	5101	CACCGCTCATGATCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5160
Db	5101	CACCGCTCATGATCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5160
QY	5161	TCACCTGCAAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCACCTTTGTGCTTGAGC	5220
Db	5161	TCACCTGCAAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCACCTTTGTGCTTGAGC	5220
QY	5221	TCCTCTCTGATCAGAGCTGAGAGGCTGAGCCGGATCTTTGAAACAGGCTCTCTCTTATCT	5280
Db	5221	TCCTCTCTGATCAGAGCTGAGAGGCTGAGCCGGATCTTTGAAACAGGCTCTCTCTTATCT	5280
QY	5281	TCCCGCATCTTCTGCTGGGCGGGGCTCATTTGACATGTTGGGAAACAGGCTGCTG	5340
Db	5281	TCCCGCATCTTCTGCTGGGCGGGGCTCATTTGACATGTTGGGAAACAGGCTGCTG	5340
QY	5341	ATGCTTTGAGCGCTTGGGAGAGAGGAGTCCAGTCCACCTGCGTGGGAGGTGCTG	5400
Db	5341	ATGCTTTGAGCGCTTGGGAGAGAGGAGTCCAGTCCACCTGCGTGGGAGGTGCTG	5400
QY	5401	GCAAGAACCTCTTGGCCATGATACAGGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5460
Db	5401	GCAAGAACCTCTTGGCCATGATACAGGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5460
QY	5461	TGAGCAGCAGAGCAGCTCTCTGCTCAGCCAGCGGCTGAGGCTCTCTCTCTCTCTCTCTCT	5520
Db	5461	TGAGCAGCAGAGCAGCTCTCTGCTCAGCCAGCGGCTGAGGCTCTCTCTCTCTCTCTCTCT	5520
QY	5521	AGGAGGACGAGGATGTAGCCGCTGAAAGGAGGCGGTGGTCCAAAGGAGCCACCGGGGG	5580

Db	5521	AGGAGGACGAGGATGTAGCCCGTGAACGGAGCGGTGGTCCAAAGAGCCACCCAGGGG	5580
QY	5581	ATGTGTTGTGTGAGGAACTTTGACCAAGGTATACCGTGGGCGAGAGATGCCAGCTGTTG	5640
Db	5581	ATGTGTTGTGTGAGGAACTTTGACCAAGGTATACCGTGGGCGAGAGATGCCAGCTGTTG	5640
QY	5641	ACCGCTTGTGCTGGGATTTCCCTCTGCTGAGTGTGTTTGGGCTGCTGGGTGTAATGGAG	5700
Db	5641	ACCGCTTGTGCTGGGATTTCCCTCTGCTGAGTGTGTTTGGGCTGCTGGGTGTAATGGAG	5700
QY	5701	CAGGAAAGACGTCACCTTTTCGATGTCAGCGGGGACACATTTGGGCGAGCAGGGGCGAGG	5760
Db	5701	CAGGAAAGACGTCACCTTTTCGATGTCAGCGGGGACACATTTGGGCGAGCAGGGGCGAGG	5760
QY	5761	CTGTGCTGGCAGGCAACAGCTGGGCGGGAACCCAGTGTCTGCGCACCTCAGCATGGAT	5820
Db	5761	CTGTGCTGGCAGGCAACAGCTGGGCGGGAACCCAGTGTCTGCGCACCTCAGCATGGAT	5820
QY	5821	ACTGCCCTCAATCCGATGCCATCTTTTGGCTGCTGACGGGCGCGAGCACCTCGAGCTGC	5880
Db	5821	ACTGCCCTCAATCCGATGCCATCTTTTGGCTGCTGACGGGCGCGAGCACCTCGAGCTGC	5880
QY	5881	TTGCGGCGCTGCGCGGTGTCCCGAGGCGCAGGTTGCTCCAGACCGCTGCTCAGGCTGG	5940
Db	5881	TTGCGGCGCTGCGCGGTGTCCCGAGGCGCAGGTTGCTCCAGACCGCTGCTCAGGCTGG	5940
QY	5941	CGGCTCTGGGACTCTCATGTACGACAGCCGGCTCAGGCACTTACAGCGAGGGAACA	6000
Db	5941	CGGCTCTGGGACTCTCATGTACGACAGCCGGCTCAGGCACTTACAGCGAGGGAACA	6000
QY	6001	MACGCAAGCTGGGAGCGGCTGGCTGTTGGGACCCAGCCGCTGCTGTTCTGAGC	6060
Db	6001	MACGCAAGCTGGGAGCGGCTGGCTGTTGGGACCCAGCCGCTGCTGTTCTGAGC	6060
QY	6061	AGCGCAACACAGGACATGAGCCCGCGGCTTCTCTTTTGGAAACAGCTTTTGGCGG	6120
Db	6061	AGCGCAACACAGGACATGAGCCCGCGGCTTCTCTTTTGGAAACAGCTTTTGGCGG	6120
QY	6121	TGTTGCGGAGGCGCTTCACTGATGCTCACTTCCATAGCATGAGAGGATGTGAAGCG	6180
Db	6121	TGTTGCGGAGGCGCTTCACTGATGCTCACTTCCATAGCATGAGAGGATGTGAAGCG	6180
QY	6181	TCCTGCTCGGCTTACGATCATGTTGAAATGGGCGGTTCCGCTGCTGGGAGCCGCAAC	6240
Db	6181	TCCTGCTCGGCTTACGATCATGTTGAAATGGGCGGTTCCGCTGCTGGGAGCCGCAAC	6240
QY	6241	ATCTCAAGGCGAGATTCGCGGCGGGTCAACACTGACCTTCCGCGGCTGCGCGCAAGGT	6300
Db	6241	ATCTCAAGGCGAGATTCGCGGCGGGTCAACACTGACCTTCCGCGGCTGCGCGCAAGGT	6300
QY	6301	CCGAGCGGCGAGCGCTTCTGTTGGCGGCGAGTTCCTTGGTTCGAGCTGCGGAGGAC	6360
Db	6301	CCGAGCGGCGAGCGCTTCTGTTGGCGGCGAGTTCCTTGGTTCGAGCTGCGGAGGAC	6360
QY	6361	ATGAGGCGCGCTGCGCTTCCAGCTGCGCGGAGGCGCTGCGGCTGCGCGGCTGCT	6420
Db	6361	ATGAGGCGCGCTGCGCTTCCAGCTGCGCGGAGGCGCTGCGGCTGCGCGGCTGCT	6420
QY	6421	TTGGAGAGCTGGCGGTGACCGCGCAGAGCACCGGCTTGGAGGACTTTTCCGTGAGCCAGA	6480
Db	6421	TTGGAGAGCTGGCGGTGACCGCGCAGAGCACCGGCTTGGAGGACTTTTCCGTGAGCCAGA	6480
QY	6481	CGATGCTGAGGAGGATTTCTTGTATTTCTTCAAGGACAGCGGGAAGGAGCAGGACCG	6540
Db	6481	CGATGCTGAGGAGGATTTCTTGTATTTCTTCAAGGACAGCGGGAAGGAGCAGGACCG	6540
QY	6541	AAGAGCAGAGGAGGAGGAGTGGGAGTGGACCCCGGCGCAGGCTGCGACACCCCAAC	6600
Db	6541	AAGAGCAGAGGAGGAGGAGTGGGAGTGGACCCCGGCGCAGGCTGCGACACCCCAAC	6600
QY	6601	GCCTCAGGCGAGTTCCTTCGATGACCTTAGCACTGCGCGAGACTGTGCTCTGAGCCTCCTCC	6660

1260	Db	 CAGGATGAAGGAAAGGACGCCAGACCTTGGAGGGCCGGGACCATGTGAGGAGCCCTCGCA	1319
1332	Qy	 TCCCTTTCTGGACCTCTGGAGCGGTGGCTTACAGCTTGGCAGGACGCACACGCTGATGTGGGG	1391
1320	Db	 TCCCTTTCTGGACCTCTGGAGCGGTGGCTTACAGCTTGGCAGGACGCACACGCTGATGTGGGG	1379
1392	Qy	 CACCTGGTGGGCACGCTTGGGCCGAGTGACGGAGTGCCTGTCTTTGGACAAAGCTTGGAGCGC	1451
1380	Db	 CACCTGGTGGGCACGCTTGGGCCGAGTGACGGAGTGCCTGTCTTTGGACAAAGCTTGGAGCGC	1439
1452	Qy	 GCACCTTGAGGACGACCTCTGGTGTTCGGGGCCCTTGCAATGTCTCGCGGAACATCGATTC	1511
1440	Db	 GCACCTTGAGGACGACCTCTGGTGTTCGGGGCCCTTGCAATGTCTCGCGGAACATCGATTC	1499
1512	Qy	 TGGGCGGGCGTCTTCTTTTGGGACCTGAGGACCTCTTCAGACCCACAGACGACCCCAACC	1571
1500	Db	 TGGGCGGGCGTCTTCTTTTGGGACCTGAGGACCTCTTCAGACCCACAGACGACCCCAACC	1559
1572	Qy	 CCAGACCTTGGGCCCCGGCCACGTCGCATCAAAATCGCATGGACATTTGACGTGGTCAAG	1631
1560	Db	 CCAGACCTTGGGCCCCGGCCACGTCGCATCAAAATCGCATGGACATTTGACGTGGTCAAG	1619
1632	Qy	 AGGACCAATGAATCAGGACAGGTTTTTGGGACCTTGGCCACGCGCGGACCCCTTGACC	1691
1620	Db	 AGGACCAATGAATCAGGACAGGTTTTTGGGACCTTGGCCACGCGCGGACCCCTTGACC	1679
1692	Qy	 GACCTGGCTACGTTGTGGGGGGCTTCGTGTACTGTCAAGACCTGTGTGGAGCGTGACGC	1751
1680	Db	 GACCTGGCTACGTTGTGGGGGGCTTCGTGTACTGTCAAGACCTGTGTGGAGCGTGACGC	1739
1752	Qy	 GTCCGCGTGTCTACGCGCGCCAAACCCCGGGCGGGCTCTACTGTCAAGACGATGCCTCAT	1811
1740	Db	 GTCCGCGTGTCTACGCGCGCCAAACCCCGGGCGGGCTCTACTGTCAAGACGATGCCTCAT	1799
1812	Qy	 CCGTGCTATGTGGACGACGTGTTCTCCGTGTGTGTGAGCCGGTGCCTTCTTCTCTG	1871
1800	Db	 CCGTGCTATGTGGACGACGTGTTCTCCGTGTGTGTGAGCCGGTGCCTTCTTCTCTG	1859
1872	Qy	 ACGCTGGCTTGGATCTACTCGGTGACACTGACAGTGAAGCCGTGTGTGCGGAGAGAGAG	1931
1860	Db	 ACGCTGGCTTGGATCTACTCGGTGACACTGACAGTGAAGCCGTGTGTGCGGAGAGAGAG	1919
1932	Qy	 ACGCGGCTGCGGGACACATGCGCGGCCATGGGGCTCAGCCGCGGGTGTCTCGCTAGGC	1991
1920	Db	 ACGCGGCTGCGGGACACATGCGCGGCCATGGGGCTCAGCCGCGGGTGTCTCGCTAGGC	1979
1992	Qy	 TGGTTCTCAGCTGCTCGGGCCCTTCTCTGCTCAGCGCCGCTGTGTGGTTCTTGGTCTC	2051
1980	Db	 TGGTTCTCAGCTGCTCGGGCCCTTCTCTGCTCAGCGCCGCTGTGTGGTTCTTGGTCTC	2039
2052	Qy	 AAGCTGGGGAGACATCTCTCCCTACAGCCACCGGGGCGTGTCTTCTGTCTTTGGGAGCC	2111
2040	Db	 AAGCTGGGGAGACATCTCTCCCTACAGCCACCGGGGCGTGTCTTCTGTCTTTGGGAGCC	2099
2112	Qy	 TTGCGGTGGCCACGGTGACCAAGACTTCTGTCTCAGCGCTTCTTCTCCCGCGCCAAC	2171
2100	Db	 TTGCGGTGGCCACGGTGACCAAGACTTCTGTCTCAGCGCTTCTTCTCCCGCGCCAAC	2159
2172	Qy	 CTGGCTCGGGCTCGGGCGGGCTTGGCCCTACTTCTCCCTCTACCTGCCCCTACGTGCTGT	2231
2160	Db	 CTGGCTCGGGCTCGGGCGGGCTTGGCCCTACTTCTCCCTCTACCTGCCCCTACGTGCTGT	2219
2232	Qy	 GTGGCTTGGGGAGACCGGCTGCCCGCGGGTGGCCGCGTGGCGGAGCTGTGTGCGCC	2291
2220	Db	 GTGGCTTGGGGAGACCGGCTGCCCGCGGGTGGCCGCGTGGCGGAGCTGTGTGCGCC	2279
2292	Qy	 GTGGCTTTCGGCTTCGCGTCGAGAGCCTTGGCTCTGCTGTGAGGAGCAGGCGGAGGGCGG	2351
2280	Db	 GTGGCTTTCGGCTTCGCGTCGAGAGCCTTGGCTCTGCTGTGAGGAGCAGGCGGAGGGCGG	2339
2352	Qy	 CAGTGGGCACAACTGTGGGCACCCCGGCTTACGGCAGACGCTCTTTCAGCCTTGGCCCAAGTCTCT	2411

Db	2340	CAGTGGCACAACGTTGGGCA	CCCCGGCCTTACGGCAGACG	CTCTTTACGCTTGGCCCCAGGTC	CTCT	2399
Qy	2412	GGCCTTCTGCTGTGAACGGCGGCTCTA	CGGCGCTCGCCACCTGGTATC	CTCTGGGAAGCTGTG		2471
Db	2400	GGCCTTCTGCTGTGNACGGCGGCTCTA	CGGCGCTCGCCACCTGGTATC	CTCTGGGAAGCTGTG		2459
Qy	2472	TGCCAGGCCAGTACGGGATCCTTGAA	CCATAGGAAATTTTCTTTT	CGGAGGAGCTACTGG		2531
Db	2460	TGCCAGGCCAGTACGGGATCCTTGAA	CCATAGGAAATTTTCTTTT	CGGAGGAGCTACTGG		2519
Qy	2532	TGGGAGCTCGGCGCCCAAGAGTCCAG	CGCCCTTGC	CCCAACCGCTG	GACCCAAAGGTC	2591
Db	2520	TGGGAGCTCGGCGCCCAAGAGTCCAG	CGCCCTTGC	CCCAACCGCTG	GACCCAAAGGTC	2579
Qy	2592	CTGTTAGAAAGGACACCGCCCGGCT	TGAGTCTTGGGATATC	CGGTTCG	CAGCCTGGAGAAG	2651
Db	2580	CTGTTAGAAAGGACACCGCCCGGCT	TGAGTCTTGGGATATC	CGGTTCG	CAGCCTGGAGAAG	2639
Qy	2652	CGCTTTCTTGGAAAGCCGACAGCCG	CTCGGGGGCTCAGCCTTG	ACTTCTAC	CAGGGC	2711
Db	2640	CGCTTTCTTGGAAAGCCGACAGCCG	CTCGGGGGCTCAGCCTTG	ACTTCTAC	CAGGGC	2699
Qy	2712	CACATCACCGCTTCTTGGGCGCA	CAACGGGGCCGGAACAC	ACACCTGTCTCATCTTG		2771
Db	2700	CACATCACCGCTTCTTGGGCGCA	CAACGGGGCCGGAACAC	ACACCTGTCTCATCTTG		2759
Qy	2772	AGTGGCCTCTTCCACACAGTGTGG	CTCTGCCCTTCACTCGGCGCC	CAGACGCTCCGCTCC		2831
Db	2760	AGTGGCCTCTTCCACACAGTGTGG	CTCTGCCCTTCACTCGGCGCC	CAGACGCTCCGCTCC		2819
Qy	2832	AGCATGGCCGCCATCTCGGCGCC	CACTCGGGCGTCTGCTCCTCAG	TACAACTGCTGTTTGAC		2891
Db	2820	AGCATGGCCGCCATCTCGGCGCC	CACTCGGGCGTCTGCTCCTCAG	TACAACTGCTGTTTGAC		2879
Qy	2892	ATGCTGACCGTGAACGAGACG	CTGGTTCTATGGGCGGCTGA	AGGGTCTGAGTGGCGCT		2951
Db	2880	ATGCTGACCGTGAACGAGACG	CTGGTTCTATGGGCGGCTGA	AGGGTCTGAGTGGCGCT		2939
Qy	2952	GTAGTGGGCCCCGAGCAGGACCG	CTGTCTGACGATGTGGGGCTGT	CTCTCAAGCAGAGT		3011
Db	2940	GTAGTGGGCCCCGAGCAGGACCG	CTGTCTGACGATGTGGGGCTGT	CTCTCAAGCAGAGT		2999
Qy	3012	GTGCAGACTCGCCACCTCTCTGT	TGGGATGC	AACGGAAAGCTGCTCCGTGGCCAT	TGCGCTTT	3071
Db	3000	GTGCAGACTCGCCACCTCTCTGT	TGGGATGC	AACGGAAAGCTGCTCCGTGGCCAT	TGCGCTTT	3059
Qy	3072	GTGGGCGGCTCCCAAGTTGTTAT	CTCTGGACGACCTACGGCTGG	CGTGTGATCTCGCTTCC		3131
Db	3060	GTGGGCGGCTCCCAAGTTGTTAT	CTCTGGACGACCTACGGCTGG	CGTGTGATCTCGCTTCC		3119
Qy	3132	CGCGCGGATTTTGGGAGCTGCT	CAATACCGGAGAGGTCG	ACGCTGATCTCTCC		3191
Db	3120	CGCGCGGATTTTGGGAGCTGCT	CAATACCGGAGAGGTCG	ACGCTGATCTCTCC		3179
Qy	3192	ACCCACCACTGGATGAGGCAGAG	CTGCTGGGAGACCGTGTGGCT	GTGTGGCAGAGTGGC		3251
Db	3180	ACCCACCACTGGATGAGGCAGAG	CTGCTGGGAGACCGTGTGGC	GTGTGGCAGAGTGGC		3239
Qy	3252	CGCTTGTGCTGTGGCTCCCACT	TTCTCGCGCGTCACTGGGCT	CCGCGTACTAC		3311
Db	3240	CGCTTGTGCTGTGGATCCCCAC	TCTTCTCGCGCGTCACTGGG	CTACTAC		3299
Qy	3312	CTGACGCTGTGAAGCCCGCTG	CCCCCTGACCAACATGAG	AGGCTGACCTGACATG		3371
Db	3300	CTGACGCTGTGAAGCCCGCTG	CCCCCTGACCAACATGAG	AGGCTGACCTGACATG		3359
Qy	3372	GAGGGCAGTGTGACACACG	CCAGGAAAGAAATGGC	AGCCAGGCGACAGTCCGC		3431
Db	3360	GAGGGCAGTGTGACACACG	CCAGGAAAGAAATGGC	AGCCAGGCGACAGTCCGC		3419
Qy	3432	ACTCCTCAGCTGTGGCCCTTG	TACAGCACTGGGTG	CCCGGGCA	CGGTGTGGAGGAG	3491
Db	3420	ACTCCTCAGCTGTGGCCCTTG	TACAGCACTGGGTG	CCCGGGCA	CGGTGTGGAGGAG	3479

Db 535 GGAAGCTGATCCCAAGCTGAGGGCTGCACGACGACGCGCCAGCCTCAACCAACCAAGC 594
Qy 646 AGTCTCCATGGAACCAACCAATGCTGGATGTGCGGAGCTGCTGACCTGCTGCGCA 705
Db 595 AGTCTCCATGGAACCAACCAATGCTGGATGTGCGGAGCTGCTGACCTGCTGCGCA 654
Qy 706 CGGAATCCCTGGGGTTGACATGCGGCCAAGCCAGGAGCCCTTGACAGCTTTGTTGGAGG 765
Db 655 CGGAATCCCTGGGGTTGACATGCGGCCAAGCCAGGAGCCCTTGACAGCTTTGTTGGAGG 714
Qy 766 CGCTGAGGACCTGCGCCAGAGCTCCCTGGGCTGCGGACGCTGCTGAGCTTTGGGAC 825
Db 715 CGCTGGGGACCTGCGCCAGAGCTCCCTGGGCTGCGGACGCTGCTGAGCTTTGGGAC 774
Qy 826 TGCTGACAGACCCCGAGGACCAAGCGCCCTCTGGAGTTGCTGTCAGAGCCCTCTGCA 885
Db 775 TGCTGACAGACCCCGAGGACCAAGCGCCCTCTGGAGTTGCTGTCAGAGCCCTCTGCA 834
Qy 886 GTGTCAGGGGACCTAGCAGCAGTGGGCGCCCTCCCTCAACTGCTGACAGGCTAGTGACC 945
Db 835 GTGTCAGGGGACCTAGCAGCAGTGGGCGCCCTCCCTCAACTGCTGACAGGCTAGTGACC 894
Qy 946 TGATGGAGCTGTGGGGCAGGAGCCAGAAATCGGCCCTGCCAGACAGCAGCCTGAGCCCG 1005
Db 895 TGATGGAGCTGTGGGGCAGGAGCCAGAAATCGGCCCTGCCAGACAGCAGCCTGAGCCCG 954
Qy 1006 CTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCGCTGTCGCCCTGCTGCGAGAC 1065
Db 955 CTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCGCTGTCGCCCTGCTGCGAGAC 1014
Qy 1066 GCCTGAAGCCTGATCTCGGGAAGCTACTCTTTGACACGATACACCTTTTACCAGGA 1125
Db 1015 GCCTGAAGCCTGATCTCGGGAAGCTACTCTTTGACACGATACACCTTTTACCAGGA 1074
Qy 1126 AGCTCATGGCCAGTGAACCGGACCTTCGAGGAGCTCACCTGCTGAGGAGTGTCCGG 1185
Db 1075 AGCTCATGGCCAGTGAACCGGACCTTCGAGGAGCTCACCTGCTGAGGAGTGTCCGG 1134
Qy 1186 AGGTGTGGAGATGTGGGACCCCGGATCTTACCTTCATGAACGACAGTTCGAATGTGG 1245
Db 1135 AGGTGTGGAGATGTGGGACCCCGGATCTTACCTTCATGAACGACAGTTCGAATGTGG 1194
Qy 1246 CCATGCTCAGCGGCTCTGACAGATGACAGGATGAAGGAAGGAGGAGCCAGACCTGGAG 1305
Db 1195 CCATGCTCAGCGGCTCTGACAGATGACAGGATGAAGGAAGGAGGAGCCAGACCTGGAG 1254
Qy 1306 GCCGGACCATGAGGACCTCGATCTCTTCTGACCTTCGACCTGCGGAGCGGTGACAGCT 1365
Db 1255 GCCGGACCATGAGGACCTCGATCTCTTCTGACCTTCGACCTGCGGAGCGGTGACAGCT 1314
Qy 1366 GCCAGGACCATGATGTGGGACCTGCTGGGACGCTGCGGCGAGTGACGGAGT 1425
Db 1315 GCCAGGACCATGATGTGGGACCTGCTGGGACGCTGCGGCGAGTGACGGAGT 1374
Qy 1426 GCCTGTCTTGGACAAGCTGAGGCGGACCCCTCAGAGGACGCTGCTGCTGCGGAGCC 1485
Db 1375 GCCTGTCTTGGACAAGCTGAGGCGGACCCCTCAGAGGACGCTGCTGCTGCGGAGCC 1434
Qy 1486 TGCAACTGCTCGCGGACATGATCTTGGGCGGCGGTGCTTCTTGGGACCTGAGACT 1545
Db 1435 TGCAACTGCTCGCGGACATGATCTTGGGCGGCGGTGCTTCTTGGGACCTGAGACT 1494
Qy 1546 CTTGAGACCCACAGACACCAACCCAGACCTGGGCGCGGACGCTGCGGACATCAAAA 1605
Db 1495 CTTGAGACCCACAGACACCAACCCAGACCTGGGCGCGGACGCTGCGGACATCAAAA 1554
Qy 1606 TCCGATGGACATTTGACGTGTGTCAGGAGCAATAAGATCAGGAGCAGGTTTGGGAC 1665
Db 1555 TCCGATGGACATTTGACGTGTGTCAGGAGCAATAAGATCAGGAGCAGGTTTGGGAC 1614
Qy 1666 CTGGCCAGCGCGGACCCCTGACCGACCTGCGCTAGCTGCGGCGGCTGCTGCTGAC 1725
Db 1615 CTGGCCAGCGCGGACCCCTGACCGACCTGCGCTAGCTGCGGCGGCTGCTGCTGAC 1674

Qy 1726 TGCAAGACCTGCTGGAGCGTGCAGCCGCTCCGCTGCTCAGCGGCGCCAAACCCCGGGCG 1785
Db 1675 TGCAAGACCTGCTGGAGCGTGCAGCCGCTCCGCTGCTCAGCGGCGCCAAACCCCGGGCG 1734
Qy 1786 GCCTCTACCTGCAGAGATGCCCTATCCGCTCTATGTGGACGAGTGTTCCTGCTGTGC 1845
Db 1735 GCCTCTACCTGCAGAGATGCCCTATCCGCTCTATGTGGACGAGTGTTCCTGCTGTGC 1794
Qy 1846 TGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
Db 1795 TGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1854
Qy 1906 TGAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1965
Db 1855 TGAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1914
Qy 1966 TCAGCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2025
Db 1915 TCAGCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
Qy 2026 GCGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2085
Db 1975 GCGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
Qy 2086 GCGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2145
Db 2035 GCGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094
Qy 2146 TCAGCGGCTTCTTCTCCGCGCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2205
Db 2095 TCAGCGGCTTCTTCTCCGCGCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2154
Qy 2206 CCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2265
Db 2155 CCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2214
Qy 2266 GCGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2325
Db 2215 GCGTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274
Qy 2326 TGCTGGAGGACGAGGCGAGGCGCGCAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2385
Db 2275 TGCTGGAGGACGAGGCGAGGCGCGCAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2334
Qy 2386 ACCTTTCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2445
Db 2335 ACCTTTCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2394
Qy 2446 TCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2505
Db 2395 TCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2454
Qy 2506 ATTTTCTTTTCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2565
Db 2455 ATTTTCTTTTCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
Qy 2566 GCGCCACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2625
Db 2515 GCGCCACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2574
Qy 2626 GCGTATCCGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2685
Db 2575 GCGTATCCGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634
Qy 2686 GGCTCAGCTGCTGCTTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2745
Db 2635 GGCTCAGCTGCTTCTTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2694
Qy 2746 GCAAGACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2805
Db 2695 GCAAGACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2754

QY 2806 TCATCTGGGCCACGACGTCCGCTCCAGCATGCGCGCCATCCCGCCCCCACTCGGGGCTCT 2865
DB |||||
2755 TCATCTGGGCCACGACGTCCGCTCCAGCATGCGCGCCATCCCGCCCCCACTCGGGGCTCT 2814
QY |||||
2866 GTCTCTAGTAAACAGTCTGTTTGAATGATGTGACCGTGGAGAGAGAGTCTGTGTTCTATG 2925
DB |||||
2815 GTCTCTAGTAAACAGTCTGTTTGAATGATGTGACCGTGGAGAGAGAGTCTGTGTTCTATG 2874
QY |||||
2926 GCGGCTGAAGGCTCTGAGTGGCGCTGAGTGGGCCCCCGAGCAGGACCGTCTCTCGCAGG 2985
DB |||||
2875 GCGGCTGAAGGCTCTGAGTGGCGCTGAGTGGGCCCCCGAGCAGGACCGTCTCTCGCAGG 2934
QY |||||
2986 ATGTGGGCTGGTCTTCAAGCAGAGTGTGAGACTGCCACCTCTCTGTGGGATGCAAC 3045
DB |||||
2935 ATGTGGGCTGGTCTTCAAGCAGAGTGTGAGACTGCCACCTCTCTGTGGGATGCAAC 2994
QY |||||
3046 GGAAGCTGTCCGTGGCCATTTGCTTGTGGGCGGCTCCCAAGTGTGTTATCCTCGAGCAGC 3105
DB |||||
2995 GGAAGCTGTCCGTGGCCATTTGCTTGTGGGCGGCTCCCAAGTGTGTTATCCTCGAGCAGC 3054
QY |||||
3106 CTACGGCTGGCGTGGATCTCTCCCGCGCGGTATTTGGGAGCTGCTCTCAATATCC 3165
DB |||||
3055 CTACGGCTGGCGTGGATCTCTCTCCCGCGCGGTATTTGGGAGCTGCTCTCAATATCC 3114
QY |||||
3166 GAGAAGCTCGACGCTGATCTCTCCACCCACCACTGGATGAGGACAGCTGCTGGAG 3225
DB |||||
3115 GAGNAGTCCGACGCTGATCTCTCCACCCACCACTGGATGAGGACAGCTGCTGGAG 3174
QY |||||
3226 ACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGTGCTGCCACTCTTCTCTGC 3285
DB |||||
3175 ACCGTGTGGCGTGGTGGCAGGTGGCGCTGTGCTGTGCTGCCACTCTTCTCTGC 3234
QY |||||
3286 GCGGTACCTGGGCTCGGCTACTACCTGACGCTGTGTGAGGCGCGGCTGCGCTGACCA 3345
DB |||||
3235 GCGGTACCTGGGCTCGGCTACTACCTGACGCTGTGTGAGGCGCGGCTGCGCTGACCA 3294
QY |||||
3346 CCAATGAGAAGGCTGACACTGACATGAGGAGCAGTGTGGACACACGAGCAGGAAAAGAGA 3405
DB |||||
3295 CCAATGAGAAGGCTGACACTGACATGAGGAGCAGTGTGGACACACGAGCAGGAAAAGAGA 3354
QY |||||
3406 ATGECAGCCAGGAGCAGAGTGGGCACTCTCAGCTGTGCTGGCCCTGGTACAGCACTGGG 3465
DB |||||
3355 ATGECAGCCAGGAGCAGAGTGGGCACTCTCAGCTGTGCTGGCCCTGGTACAGCACTGGG 3414
QY |||||
3466 TGCCCGGGGACGCGCTGGTGGAGAGCTGCGCACACTCTTCCGAGAGCTAGACACGCGGCTGGCGG 3525
DB |||||
3415 TGCCCGGGGACGCGCTGGTGGAGAGCTGCGCACACTCTTCCGAGCTGGTGGTCTGCCCTACA 3474
QY |||||
3526 CCGGTGCCATGACGGGAGCTTCCGCACACTCTTCCGAGAGCTAGACACGCGGCTGGCGG 3585
DB |||||
3475 CCGGTGCCATGACGGGAGCTTCCGCACACTCTTCCGAGAGCTAGACACGCGGCTGGCGG 3534
QY |||||
3586 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACCGACCTCGAGGAGATCTTCTGAAGG 3645
DB |||||
3535 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACCGACCTCGAGGAGATCTTCTGAAGG 3594
QY |||||
3646 TGGTGGAGGAGTGTGCTGGGACACAGATATGAGGAGTGCAGCTGGGCGAGCACTAT 3705
DB |||||
3595 TGGTGGAGGAGTGTGCTGGGACACAGATATGAGGAGTGCAGCTGGGCGAGCACTAT 3654
QY |||||
3706 GCACAGGCACTTGTGGCTTAGACGTAACCTCGGGCTCAAGATGCGGCCACAGAGACAG 3765
DB |||||
3655 GCACAGGCACTTGTGGCTTAGACGTAACCTCGGGCTCAAGATGCGGCCACAGAGACAG 3714
QY |||||
3766 CGCTGGAGAACCGGGAAACAGCTGGGTGAGCCCCAGAGACTGACCGGGCTCTGGGCCAG 3825
DB |||||
3715 CGCTGGAGAACCGGGAAACAGCTGGGTGAGCCCCAGAGACTGACCGGGCTCTGGGCCAG 3774
QY |||||
3826 ACCCGTGGGCGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTGCTTC 3885
DB |||||
3775 ACCCGTGGGCGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTGCTTC 3834
QY |||||
3886 TCAAGCGCTTCTGCTTGGCCCGCGCAGCGCGCGGCTGTTCCGCCAGATCGTGTGTC 3945

DB |||||
3835 TCAAGCGCTTCTGCTTGGCCCGCGCAGCGCGCGGCTGTTCCGCCAGATCGTGTGC 3894
QY |||||
3946 CTGCCCTCTTTGTGGGCTGGGCTCGTGTTCAGGCTCTCATCGTGCCTCTCTTTGGGCACT 4005
DB |||||
3895 CTGCCCTCTTTGTGGGCTGGGCTCGTGTTCAGGCTCTCATCGTGCCTCTCTTTGGGCACT 3954
QY |||||
4006 ACCCGCTCTGGGCTCAGTCCACCACATGTACGGTGTCTCAGGTGTCTTCTTCAGTGAGG 4065
DB |||||
3955 ACCCGCTCTGGGCTCAGTCCACCACATGTACGGTGTCTCAGGTGTCTTCTTCAGTGAGG 4014
QY |||||
4066 AGCCCCAGGGGACCTTGGACGTGCTCGAGCGCTGCTCGAGCGCTGTCTGACAGGACAGGAC 4125
DB |||||
4015 AGCCCCAGGGGACCTTGGACGTGCTCGAGCGCTGCTCGAGCGCTGTCTGACAGGACAGGAC 4074
QY |||||
4126 TGGAGAGAGCCCGCAGTGCAGCATAGCTCCACAGAGTTCTCGGACACAGAGTTCTCTGCTG 4185
DB |||||
4075 TGGAGAGAGCCCGCAGTGCAGCATAGCTCCACAGAGTTCTCGGACACAGAGTTCTCTGCTG 4134
QY |||||
4186 AGTGTAGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCACAGAGTCTCCATCCCCAGGCTGCC 4245
DB |||||
4135 AGTGTAGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCACAGAGTCTCCATCCCCAGGCTGCC 4194
QY |||||
4246 AGTGTAGCCAGCCCGGTGCTGCTCGGCGCTGCTGCTCGGCTGCTGAGCTGTGCTGCT 4305
DB |||||
4195 AGTGTAGCCAGCCCGGTGCTGCTCGGCGCTGCTGCTCGGCTGCTGAGCTGTGCTGCT 4254
QY |||||
4306 CCGCTCGGCGCCAGGAGTGAACCGGCTCTGGGGAGTGGTTGAGAACCTGACAGGCGGGA 4365
DB |||||
4255 CCGCTCGGCGCCAGGAGTGAACCGGCTCTGGGGAGTGGTTGAGAACCTGACAGGCGGGA 4314
QY |||||
4366 ACCGTGTGACTTCTGCTCAAGACTACCGCGCTGTGCGCCAGGCGCTGAGGACTA 4425
DB |||||
4315 ACCGTGTGACTTCTGCTCAAGACTACCGCGCTGTGCGCCAGGCGCTGAGGACTA 4374
QY |||||
4426 AGAAGTGGGTGATGAGGTACGAGTACGAGGCTTCTCGCTGGGGGCGCCAGAGCCAGGCC 4485
DB |||||
4375 AGAAGTGGGTGATGAGGTACGAGTACGAGGCTTCTCGCTGGGGGCGCCAGAGCCAGGCC 4434
QY |||||
4486 TGGCTCGGCGCCAGAGTGGGCGGCTCAGTGGAGAGTGTGGGCGCTGCTGAGTCCCC 4545
DB |||||
4435 TGGCTCGGCGCCAGAGTGGGCGGCTCAGTGGAGAGTGTGGGCGCTGCTGAGTCCCC 4494
QY |||||
4546 TGGCTCGGCGGCGCTCGACCGTCTGAAACCTCAGAGCTGGGCTCAGAGCTGG 4605
DB |||||
4495 TGGCTCGGCGGCGCTCGACCGTCTGAAACCTCAGAGCTGGGCTCAGAGCTGG 4554
QY |||||
4606 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTGGCCT 4665
DB |||||
4555 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTGGCCT 4614
QY |||||
4666 TTGTCAACCGAGCAGCAACGCACTCTCGTGTCTACCTGCGCCCGAGGCGCGGCGCC 4725
DB |||||
4615 TTGTCAACCGAGCAGCAACGCACTCTCGTGTCTACCTGCGCCCGAGGCGCGGCGCC 4674
QY |||||
4726 AGCCCCACAGCATCACCACTCAACCCCTTGAACCTCACCAGGAGCAGCTGTCTG 4785
DB |||||
4675 ACCCCCACAGCATCACCACTCAACCCCTTGAACCTCACCAGGAGCAGCTGTCTG 4734
QY |||||
4786 AGGCTGCACGTATGGGCTCTCGTGGGAGCTCTCGTCTCTCAATCTGTGTGGTCTTTGCCA 4845
DB |||||
4735 AGGCTGCACGTATGGGCTCTCGTGGGAGCTCTCGTCTCTCAATCTGTGTGGTCTTTGCCA 4794
QY |||||
4846 TGTCTTTGTCCGGCGCAGTTCACCTCTGTCTCTCATTTGAGGAGCAGTCAACCGGAGCCA 4905
DB |||||
4795 TGTCTTTGTCCGGCGCAGTTCACCTCTGTCTCTCATTTGAGGAGCAGTCAACCGGAGCCA 4854
QY |||||
4906 AGCACCTGAGCTCATGGGCGGCTGCTCCCGACCCCTTACTGCTTGGCACTTTCTCT 4965
DB |||||
4855 AGCACCTGAGCTCATGGGCGGCTGCTCCCGACCCCTTACTGCTTGGCACTTTCTCT 4914
QY |||||
4966 GGGACATGTGTAACCTATTGGTGCCAGCATGCACTGTGTGCTCATCTTTCTGCGCTTCC 5025

4306 CCCTCCGCCCCAGGAGTACCGGCTCTGGGAAGTGGTTCAAACTCAGACGGCCGGA 4365
4255 CCCTCCGCCCCAGGAGTACCGGCTCTGGGAAGTGGTTCAAACTCAGACGGCCGGA 4314
4366 ACCTGTCTGACTTCTCTGGTCAAGACTACCGGCTCTGGTGCAGGAGGCTCAAGACTA 4425
4315 ACCTGTCTGACTTCTCTGGTCAAGACTACCGGCTCTGGTGCAGGAGGCTCAAGACTA 4374
4426 AGAAGTGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCGAGGCC 4485
4375 AGAAGTGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGGGGGCCGAGACCGAGGCC 4434
4486 TGCCCTCGGGCCCAAGAGTGTGGCCGCTCAGTGGAGGAGTGTGGCGCTCCTCAGTCCCC 4545
4435 TGCCCTCGGGCCCAAGAGTGTGGCCGCTCAGTGGAGGAGTGTGGCGCTCCTCAGTCCCC 4494
4546 TGCCCTCGGGGGCCCTCGACCGGTCTCTGAAACCTCAGAGCTGGGCTCAGAGCCTGG 4605
4495 TGCCCTCGGGGGCCCTCGACCGGTCTCTGAAACCTCAGAGCTGGGCTCAGAGCCTGG 4554
4606 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGCTGGCACTCCATGGTGGCCT 4665
4555 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGCTGGCACTCCATGGTGGCCT 4614
4666 TTGTCACCGAGCAGCAACGCAATCTCCGTGCTCAGCTGCCCCAGAGCCCGCCCGCC 4725
4615 TTGTCACCGAGCAGCAACGCAATCTCCGTGCTCAGCTGCCCCAGAGCCCGCCCGCC 4674
4726 ACGCCACAGCATCACACACTCAACCAACCCCTTGAACCTCAGCAAGGAGCAGTGTCTG 4785
4675 ACGCCACAGCATCACACACTCAACCAACCCCTTGAACCTCAGCAAGGAGCAGTGTCTG 4734
4786 AGGCTGACATGATGGCTCTCCGTGGAAGTCTCTGCTCCTCATCTGTGTGGTCTTTGCCA 4845
4735 AGGCTGACATGATGGCTCTCCGTGGAAGTCTCTGCTCCTCATCTGTGTGGTCTTTGCCA 4794
4846 TGTCTCTTGTCCGGCCAGCTTCACTCTGCTGCTCATTTGAGGAGCAGATCACCCGAGCCA 4905
4795 TGTCTCTTGTCCGGCCAGCTTCACTCTGCTGCTCATTTGAGGAGCAGATCACCCGAGCCA 4854
4906 AGCACCTTGACAGCTCATGGGGGGCTGTCCCCCACCCTCTACTGGCTTTGGCAACTTCTCT 4965
4855 AGCACCTTGACAGCTCATGGGGGGCTGTCCCCCACCCTCTACTGGCTTTGGCAACTTCTCT 4914
4966 GGGACATGTGTAACTACTTGGTGCAGCATGCAATGCTGTGTGTGTCTATTTCTGGCCTTCC 5025
4915 GGGACATGTGTAACTACTTGGTGCAGCATGCAATGCTGTGTGTGTCTATTTCTGGCCTTCC 4974
5026 AGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCTCTGCTGTGTGTGTGTGTGTGT 5085
4975 AGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCTCTGCTGTGTGTGTGTGTGTGT 5034
5086 ATGGCTGTGCTGATCACACCGCTCATGTACCCAGCTCTTCTTCTCTCGTGCAGCA 5145
5035 ATGGCTGTGCTGATCACACCGCTCATGTACCCAGCTCTTCTTCTCTCGTGCAGCA 5094
5146 CAGCCTATGTGTGCTCACTGATCAAACTCTTTATTTGATCAATTTGATCAATTTGATCA 5205
5095 CAGCCTATGTGTGCTCACTGATCAAACTCTTTATTTGATCAATTTGATCAATTTGATCA 5154
5206 CTTTGTGCTGTGAGCTCTTCTGATCAGAGCTGAGAGGAGTGGCCGATCTTGAAC 5265
5155 CTTTGTGCTGTGAGCTCTTCTGATCAGAGCTGAGAGGAGTGGCCGATCTTGAAC 5214
5266 AGGTCTTCTTATCTTCCCACTTCTGCTGTGGCCGGGGCTCAATTGATGATGGTGGGA 5325
5215 AGGTCTTCTTATCTTCCCACTTCTGCTGTGGCCGGGGCTCAATTGATGATGGTGGGA 5274
5326 ACCAGGCGATGGCTGATGCTTTGAGCGCTTTGAGAGCAGGAGTTCAGTCAACCCCTGC 5385
5275 ACCAGGCGATGGCTGATGCTTTGAGCGCTTTGAGAGCAGGAGTTCAGTCAACCCCTGC 5334
5386 GCTGGGAGGTGGTGGGCAAGAACTCTTGGCCATGGTGATACAGGGGGCCCTCTTCTTCTC 5445

5335 GCTGGAGAGTGGTGGCAAGAACTCTTGGCCATGTTGATACAGGGGCCCTCTTCTTCTC 5394
5446 TCTTCACTACTTGTCTGAGACCCGAAGCAACTCTCTGCCACAGCCAGGGTGAAGTCTC 5505
5395 TCTTCACTACTTGTCTGAGACCCGAAGCAACTCTCTGCCACAGCCAGGGTGAAGTCTC 5454
5506 TGCCACTCTCTGGAGAGGAGGAGGATGTAGCCGCTGNAACGGGAGCGGTTGGTCCAAG 5565
5455 TGCCACTCTCTGGAGAGGAGGAGGATGTAGCCGCTGNAACGGGAGCGGTTGGTCCAAG 5514
5566 GAGCCACCCAGGGGAGTGTGTGCTGTAGAGAACTTGAACAAAGGTATACCGTGGGCGAGA 5625
5515 GAGCCACCCAGGGGAGTGTGTGCTGTAGAGAACTTGAACAAAGGTATACCGTGGGCGAGA 5574
5626 GATGCCAGCTGTGTGACCCGCTTGTCTGGGAGTTCCTCTGTGTAGTGTGTGTGGGTGC 5685
5575 GATGCCAGCTGTGTGACCCGCTTGTCTGGGAGTTCCTCTGTGTAGTGTGTGTGGGTGC 5634
5686 TGGGTGTGATGAGCAGGGAGAGAGCTTCCACGTTTTCGATGTTGACGGGGGACACATTGG 5745
5635 TGGGTGTGATGAGCAGGGAGAGAGCTTCCACGTTTTCGATGTTGACGGGGGACACATTGG 5694
5746 CCAGCAGGGGCGAGGCTGTGTGCGCAGGCCACAGCGTGGCCCGGGAACCCAGTGTCTGCGC 5805
5695 CCAGCAGGGGCGAGGCTGTGTGCGCAGGCCACAGCGTGGCCCGGGAACCCAGTGTCTGCGC 5754
5806 ACCTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCGCG 5865
5755 ACCTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCGCG 5814
5866 AGCACCTTGAGCTGTGTGCGGCTTCCGGGCTTCCGGAGGCGCAGTGTGCCAGACCG 5925
5815 AGCACCTTGAGCTGTGTGCGGCTTCCGGGCTTCCGGAGGCGCAGTGTGCCAGACCG 5874
5926 CTGGCTCAGGCTTGGCGCTGTGGGACTCTCATGTACGACAGACCGGCTTGCAGGACACT 5985
5875 CTGGCTCAGGCTTGGCGCTGTGGGACTCTCATGTACGACAGACCGGCTTGCAGGACACT 5934
5986 ACAGCGAGGGGAAACAAACGCAAGCTGGCGCTTGGCGCTTGGTGTGGGAACCAAGCCG 6045
5935 ACAGCGAGGGGAAACAAACGCAAGCTGGCGCTTGGCGCTTGGTGTGGGAACCAAGCCG 5994
6046 TGTGTCTTGTGACAGCGCCGACACAGGCGATGACCCGACGCGCGCGCTTCTTCTTGA 6105
5995 TGTGTCTTGTGACAGCGCCGACACAGGCGATGACCCGACGCGCGCGCTTCTTCTTGA 6054
6106 ACAGCCTTTTGGCGCTTGGCGGAGGGCGCTTCACTGATGCTCACTCCCATAGCATGG 6165
6055 ACAGCCTTTTGGCGCTTGGCGGAGGGCGCTTCACTGATGCTCACTCCCATAGCATGG 6114
6166 AGGAGTGTGAAGCGCTCTGCTCGGCTTAGCCATCATGTGAAATGGCGGTTCCGCTGCC 6225
6115 AGGAGTGTGAAGCGCTCTGCTCGGCTTAGCCATCATGTGAAATGGCGGTTCCGCTGCC 6174
6226 TGGGAGCGCGGCAACTCTCAGGGGAGATTCGCGCGGGTCAACACTGACCTTCCGCTGCCG 6285
6175 TGGGAGCGCGGCAACTCTCAGGGGAGATTCGCGCGGGTCAACACTGACCTTCCGCTGCCG 6234
6286 TGCCCGCGCAGAGTCCAGCGGAGCGGCTTCTGGCGGCGGAGTTCCTTGGGTGCCG 6345
6235 TGCCCGCGCAGAGTCCAGCGGAGCGGCTTCTGGCGGCGGAGTTCCTTGGGTGCCG 6294
6346 AGCTGCGGAGGCACTAGGAGCGGCTTCCAGTTCAGCTGCGCGCGGAGGCGCTGCG 6405
6295 AGCTGCGGAGGCACTAGGAGCGGCTTCCAGTTCAGCTGCGCGCGGAGGCGCTGCG 6354
6406 CCCTGCGCGGCTTCTTGGAGAGCTGCGGCTGACCGCGGAGCAGCAGCGGCTGGAGACT 6465
6355 CCCTGCGCGGCTTCTTGGAGAGCTGCGGCTGACCGCGGAGCAGCAGCGGCTGGAGACT 6414
6466 TTTTCCGTGAGCCAGAGCTGCGGAGGAGTATTTCTTGTACTTCTTCAAGAGCAGCGGGA 6525

2374	DB	ACACGCTGATGTGGGGCACTGTGTGGGCA	CGCTGGGCGGAGTGCAGGAGTGCCTGTCTT	2433
1436	QY	GGACAAGCTGGAGCGGCACCTCAGAGG	CAGCCCTGGTGTGCGGGGCCCTGCAACTGCT	1495
2434	DB	GGACAAGCTGGAGCGGCACCTCAGAGG	CAGCCCTGGTGTGCGGGGCCCTGCAACTGCT	2493
1496	QY	CGCGAATATCGATTCTGGGCGGGCGTCT	TTTGGGACCTTGAGACCTCTTCAAGACC	1555
2494	DB	CGCGAATATCGATTCTGGGCGGGCGTCT	TTTGGGACCTTGAGACCTCTTCAAGACC	2553
1556	QY	CACAGACCCCAACCCAGACCTGGCCCG	CGCCAGCTGGCATCAAAATCCGATGGA	1615
2554	DB	CACAGACCCCAACCCAGACCTGGCCCG	CGCCAGCTGGCATCAAAATCCGATGGA	2613
1616	QY	CATTGACGTGGTCAACAGGACCAATAGA	TACAGGACAGCTTTTGGACCCCTGGCC	1675
2614	DB	CATTGACGTGGTCAACAGGACCAATAGA	TACAGGACAGCTTTTGGACCCCTGGCC	2673
1676	QY	CGCGGACCCCTGACACGACTCGGCTAC	GTGTGGGGCGGTTCTGTTACCTGAAGACCT	1735
2674	DB	CGCGGACCCCTGACACGACTCGGCTAC	GTGTGGGGCGGTTCTGTTACCTGAAGACCT	2733
1736	QY	GSTGGAGCGTGCAGCGCTCGCGTGTCA	CGGGGCGCAACCCCGGGCGGCTCTACCT	1795
2734	DB	GSTGGAGCGTGCAGCGCTCGCGTGTCA	CGGGGCGCAACCCCGGGCGGCTCTACCT	2793
1796	QY	GCAGCAGATGCCCTATCCGTGTATGTG	ACAGCTGTCTCGGTGTCTGAGCCGGTC	1855
2794	DB	GCAGCAGATGCCCTATCCGTGTATGTG	ACAGCTGTCTCGGTGTCTGAGCCGGTC	2853
1856	QY	GCTGCCGCTCTTCTGTACGCTGGCCT	TGGATCTACTCCGTGACACTGACGTGAAGCCGT	1915
2854	DB	GCTGCCGCTCTTCTGTACGCTGGCCT	TGGATCTACTCCGTGACACTGACGTGAAGCCGT	2913
1916	QY	GGTGCGGGAGAGAGACGGCGCTGCGG	AGACCAATGCGCGGCTGAGCGCGC	1975
2914	DB	GGTGCGGGAGAGAGACGGCGCTGCGG	AGACCAATGCGCGGCTGAGCGCGC	2973
1976	QY	GSTGCTCTGCTAGGCTGTCTCAGCTGC	CTCGGCCCTTCTGCTCAGCGCCGCGCT	2035
2974	DB	GSTGCTCTGCTAGGCTGTCTCAGCTGC	CTCGGCCCTTCTGCTCAGCGCCGCGCT	3033
2036	QY	GCTGGTTCTGGTGTCAAGCTGGGGG	ACATCTCTCCCTACAGCCACCCGGGCGTGTCT	2095
3034	DB	GCTGGTTCTGGTGTCAAGCTGGGGG	ACATCTCTCCCTACAGCCACCCGGGCGTGTCT	3093
2096	QY	CTGTCTTGGGACGCTTCCGGTGGGCA	CGGTGACCAAGCTTCTGTCTGAGCGGCTT	2155
3094	DB	CTGTCTTGGGACGCTTCCGGTGGGCA	CGGTGACCAAGCTTCTGTCTGAGCGGCTT	3153
2156	QY	CTTCTCCGCGCAACTGTGCTCGGCT	CGGCTGCGCGGCTGCTCTCTCTCTACCT	2215
3154	DB	CTTCTCCGCGCAACTGTGCTCGGCT	CGGCTGCGCGGCTGCTCTCTCTCTACCT	3213
2216	QY	GCCTACGTGCTGTGTGCTTGGCGG	ACCGGCTGCGCGGCTGCGCGCGC	2275
3214	DB	GCCTACGTGCTGTGTGCTTGGCGG	ACCGGCTGCGCGGCTGCGCGCGC	3273
2276	QY	GAGCTGCTGTGCGCCGTGGCGCTT	CGGCTTGGCATGAGCCTGTGTGAGGA	2335
3274	DB	GAGCTGCTGTGCGCCGTGGCGCTT	CGGCTTGGCATGAGCCTGTGTGAGGA	3333
2336	QY	GCAGGGCGGGCGCAGTGGACACAGT	GGGACCCGGCTTACGGCACGCTTTCAG	2395
3334	DB	GCAGGGCGGGCGCAGTGGACACAGT	GGGACCCGGCTTACGGCACGCTTTCAG	3393
2396	QY	CTTGGCCAGGCTCTCTGGCCCTTCT	GTGTGTGAGACGGCGGCTCTACGGCCCTGCCACCTG	2455
3394	DB	CTTGGCCAGGCTCTCTGGCCCTTCT	GTGTGTGAGACGGCGGCTCTACGGCCCTGCCACCTG	3453
2456	QY	GTAACCTGGAGCTGTGTGCCAGCC	AGTACGGGATCCCTGAAACCATGGAATTTCTTTT	2515
3454	DB	GTAACCTGGAGCTGTGTGCCAGCC	AGTACGGGATCCCTGAAACCATGGAATTTCTTTT	3513

Qy	2516	TCGGAGAGCTACTGTGGGACCTCGGGCCCCCAAGAGTFCAGCCCTTGCCTACCCC	2575
Db	3514	TCGGAGAGCTACTGTGGGACCTCGGGCCCCCAAGAGTCCAGCCCTTGCCCAACCC	3573
Qy	2576	GCTGGACCCAAAGGTGCTGTAGAAGGACCGCCGCCCTGAGTCTTGCGCTACCGT	2635
Db	3574	GCTGGACCCAAAGGTGCTGTAGAAGGACCGCCGCCCTGAGTCTTGCGCTTCCGT	3633
Qy	2636	TCGCAGCCTGGGAAGCGCTTTCCTGGAAAGCCCGCAGCCAGCCCTCGGGGGCTCAGCCT	2695
Db	3634	TCGCAGCCTGGGAAGCGCTTTCCTGGAAAGCCCGCAGCCAGCCCTCGGGGGCTCAGCCT	3693
Qy	2696	GGACTTCTACCAAGGCCACATCAACCGCTTCTTGGGCGACAACGGGGCCGGCAAGACCAC	2755
Db	3694	GGACTTCTACCAAGGGCCACATCAACCGCTTCTTGGGCGACAACGGGGCCGGCAAGACCAC	3753
Qy	2756	CACCTGTCCATCTTGAGTGGCCCTTCCACCACCAAGTGGTCTGTCCTTCACTCTGGG	2815
Db	3754	CACCTGTCCATCTTGAGTGGCCCTTCCACCACCAAGTGGTCTGTCCTTCACTCTGGG	3813
Qy	2816	CCACGAGCTCCGCTCCAGCATGCGCGCCATCCCGCCCCACCTGGGGCTGTCTCTCAGTA	2875
Db	3814	CCACGAGCTCCGCTCCAGCATGCGCGCCATCCCGCCCCACCTGGGGCTGTCTCTCAGTA	3873
Qy	2876	CAACGTGCTGTTTGAACATGCTGACCGTGAAGAGACGTCGTCTATGGGCGGCTGAA	2935
Db	3874	CAACGTGCTGTTTGAACATGCTGACCGTGAAGAGACGTCGTCTATGGGCGGCTGAA	3933
Qy	2936	GGGTCTGAGTGCCTGTAGTGGGCCCCGAGCAGGACCGTCTGCTCAGAGATCTGGGGCT	2995
Db	3934	GGGTCTGAGTGCCTGTAGTGGGCCCCGAGCAGGACCGTCTGCTCAGAGATCTGGGGCT	3993
Qy	2996	GGTCTTCAAGCAGAGTGTGAGACTCGCCACCTCTCTGGTGGGATCAACGGGAAGCTGC	3055
Db	3994	GGTCTTCAAGCAGAGTGTGAGACTCGCCACCTCTCTGGTGGGATCAACGGGAAGCTGC	4053
Qy	3056	CGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTTGTTATCTGGAGCAGGCTACGGCTGG	3115
Db	4054	CGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTTGTTATCTGGAGCAGGCTACGGCTGG	4113
Qy	3116	CGTGGATCTGCTTCCGCGCGGCTATTTGGGAGCTGCTGCTCAATACCGAAGAGTGC	3175
Db	4114	CGTGGATCTGCTTCCGCGCGGCTATTTGGGAGCTGCTGCTCAATACCGAAGAGTGC	4173
Qy	3176	CACGCTGATCCTTCCACCCACCACTGGATGAGGAGAGCTGCTGGGAGACGCTGTGGC	3235
Db	4174	CACGCTGATCCTTCCACCCACCACTGGATGAGGAGAGCTGCTGGGAGACGCTGTGGC	4233
Qy	3236	TGTGGTGGCAGTGGCGCTTGTGCTGTGCTGCCCTCCCACTCTTCTCGCGCGTCACT	3295
Db	4234	CGTGGTGGCAGTGGCGCTTGTGCTGTGCTGCCCTCCCACTCTTCTCGCGCGTCACT	4293
Qy	3296	GGGCTCCGGCTACTACTGACGCTGTGAAGGCCCGCTGCCCTGCACACCAATGAGAA	3355
Db	4294	GGGCTCCGGCTACTACTGACGCTGTGAAGGCCCGCTGCCCTGCACACCAATGAGAA	4353
Qy	3356	GGCTGACACTGACATGAGAGGCGAGTGTGACACCAAGCAGGAAAGAAATGGCAGCCA	3415
Db	4354	GGCTGACACTGACATGAGAGGCGAGTGTGACACCAAGCAGGAAAGAAATGGCAGCCA	4413
Qy	3416	GGGAGCAGAGTGGGCATCTCAGTGTCTGGGCCCTGGTACAGACTGTGGTGGCCGGGGC	3475
Db	4414	GGGAGCAGAGTGGGCATCTCAGTGTCTGGGCCCTGGTACAGACTGTGGTGGCCGGGGC	4473
Qy	3476	ACGGCTGGTGGAGAGCTGCGCACACGAGCTGTGTGCTGCTGCCCTTACACGGGTGCCCA	3535
Db	4474	ACGGCTGGTGGAGAGCTGCGCACACGAGCTGTGTGCTGCTGCCCTTACACGGGTGCCCA	4533
Qy	3536	TGACGGCAGCTTTCGCCACACTCTTCCAGAGCTTAGACACCGCGCTGGCGGAGCTGAGGCT	3595
Db	4534	TGACGGCAGCTTTCGCCACACTCTTCCAGAGCTTAGACACCGCGCTGGCGGAGCTGAGGCT	4593

3596 QY CACTGGCTACGGGATCTCGGACACAGAGCTCGAGAGATCTTCTGAAGGTGGTGAGGA 3655
4594 Db CACTGGCTACGGGATCTCGGACACAGAGCTCGAGAGATCTTCTGAAGGTGGTGAGGA 4653
3656 QY GTGTGCTCGGACACAGATATGGAGATGGAGCTGCGGGACAGACCTATGACAGGCAT 3715
4654 Db GTGTGCTCGGACACAGATATGGAGATGGAGCTGCGGGACAGACCTATGACAGGCAT 4713
3716 QY TGCTGGCTTAGAGTAACCTTCGGCTCAAGATGCGGCCACAGGAGACAGCGCTGGAGAA 3775
4714 Db TGCTGGCTTAGAGTAACCTTCGGCTCAAGATGCGGCCACAGGAGACAGCGCTGGAGAA 4773
3776 QY CGGGAAACAGCTGGGTGAGCCCAAGACATGACAGAGCTCTGGGCGAGACGCGGTGGG 3835
4774 Db CGGGAAACAGCTGGGTGAGCCCAAGACATGACAGAGCTCTGGGCGAGACGCGGTGGG 4833
3836 QY CGGGTACAGGCTGGGCACTGACCCGACAGCTCCAGGCGCTTCTCAAGCGCTT 3895
4834 Db CGGGTACAGGCTGGGCACTGACCCGACAGCTCCAGGCGCTTCTCAAGCGCTT 4893
3896 QY TGTGCTTGGCCCGCAGCGCGCGGCTGTTCGCCAGATCGTGTGCTGCGCTTCTT 3955
4894 Db TGTGCTTGGCCCGCAGCGCGCGGCTGTTCGCCAGATCGTGTGCTGCGCTTCTT 4953
3956 QY TGTGGGCTTGGCTTGTTCAGCTCATGTGTGCTCTTTCGGGCACTACCGGCTCT 4015
4954 Db TGTGGGCTTGGCTTGTTCAGCTCATGTGTGCTCTTTCGGGCACTACCGGCTCT 5013
4016 QY GCGGCTCAGTCCACCATGTAGGCTCAGGTGCTTCTTCAGTAGGAGCGCCACGG 4075
5014 Db GCGGCTCAGTCCACCATGTAGGCTCAGGTGCTTCTTCAGTAGGAGCGCCACGG 5073
4076 QY GGACCTTGGAGGTGCGCGGCTGCTCGAGGCGTGTGCGAGGAGGAGCTGGAGAGCC 4135
5074 Db GGACCTTGGAGGTGCGCGGCTGCTCGAGGCGTGTGCGAGGAGGAGCTGGAGAGCC 5133
4136 QY CCCAGTCAGCATAGCTCCACAGGTTCTCGGACACAGAGTTCTGCTGAAGTGGCCAA 4195
5134 Db CCCAGTCAGCATAGCTCCACAGGTTCTCGGACACAGAGTTCTGCTGAAGTGGCCAA 5193
4196 QY GGTCTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCGAGTGTAGCCA 4255
5194 Db GGTCTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCGAGTGTAGCCG 5253
4256 QY GCCCGTTCGCGGCGCTGCTGCCGAGTGTGCCGCTGCGAGTGTGTGCTCCCTCGCC 4315
5254 Db GCCCGTTCGCGGCGCTGCTGCCGAGTGTGCCGCTGCGAGTGTGTGCTCCCTCGCC 5313
4316 QY CCAGCATGACCGGCTCTGGGAGAGTGGTTACAGAACCTGACAGCCCGGAACTGTCTGA 4375
5314 Db CCAGCATGACCGGCTCTGGGAGAGTGGTTACAGAACCTGACAGCCCGGAACTGTCTGA 5373
4376 QY CTTCTGTGTCAAGACTACCGCGCTTGGTGGCCAGGCGCTGAAGACTAAGAGTGGGT 4435
5374 Db CTTCTGTGTCAAGACTACCGCGCGCTTGGTGGCCAGGCGCTGAAGACTAAGAGTGGGT 5433
4436 QY GAATGAGTACAGGTACGAGGCTTCTCGCTGGGGGCGAGACCCAGGCGCTGCCCTCGG 4495
5434 Db GAATGAGTACAGGTACGAGGCTTCTCGCTGGGGGCGCGAGACCCAGGCGCTGCCCTCGG 5493
4496 QY CCAAGAGTTGGGCGCTCAGTGGAGGAGTGTGGCGCTGTGATGCCCTTGTCAACCG 4555
5494 Db CCAAGAGTTGGGCGCTCAGTGGAGGAGTGTGGCGCTGTGATGCCCTTGTCAACCG 5553
4556 QY GGCCCTCGACCGTGTCTGAACCTTCAAGCTTGGGCTCAGAGCTTGGATGTCTCAGGA 4615
5554 Db GGCCCTCGACCGTGTCTGAACCTTCAAGCTTGGGCTCAGAGCTTGGATGTCTCAGGA 5613
4616 QY CAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTGGCTTGTCAACCG 4675
5614 Db CAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTGGCTTGTCAACCG 5673
4676 QY AGCCAGCAACGCAATCTCTCGTGTCTCACTGCCCCCGAGGCGCCCGCCACGCCCCACG 4735

5674 Db AGCCAGCAACGCAATCTCTCGTGTCTCACTGCCCCAGGCGCCGCGCCACGAG 5733
4736 QY CATCACCACTCAACCAACCCCTTGAACCTCACCAAGGAGAGCTGTCTGAGGCTGCAT 4795
5734 Db CATCACCACTCAACCAACCCCTTGAACCTCACCAAGGAGAGCTGTCTGAGGCTGCAT 5793
4796 QY GATGGCTCTCTCGGTGGAGCTCTCGTCTCATCTGTGTGTCTTTGGCATGCTCTTGT 4855
5794 Db GATGGCTCTCTCGGTGGAGCTCTCGTCTCATCTGTGTGTCTTTGGCATGCTCTTGT 5853
4856 QY CCGGCGAGCTTCACTCTTGTCTCATTTGAGAGGAGCTCACCCGAGGCAAGCACTGCA 4915
5854 Db CCGGCGAGCTTCACTCTTGTCTCATTTGAGAGGAGCTCACCCGAGGCAAGCACTGCA 5913
4916 QY GCTCATGGGGGCGCTGTCTCCGCCACCTCTACTTGGCTTGGCAACTTCTCTGGGACATGT 4975
5914 Db GCTCATGGGGGCGCTGTCTCCGCCACCTCTACTTGGCTTGGCAACTTCTCTGGGACATGT 5973
4976 QY TAACTACTTGTGCGCAGCATGCTGTGTCTCATCTTCTGGGCTTCCAGCAGAGGCG 5035
5974 Db TAACTACTTGTGCGCAGCATGCTGTGTCTCATCTTCTGGGCTTCCAGCAGAGGCG 6033
5036 QY ATATGTGGCCCCGTGGCAACCTGCTCTCTGCTGTGTGTACTACTGTATGGCTGTCT 5095
6034 Db ATATGTGGCCCCGTGGCAACCTGCTCTCTGCTGTGTGTACTACTGTATGGCTGTCT 6093
5096 QY GATCACCGCTCATGTACCCAGCTCTTCTTCTCGTGGCCAGCAGACCTATGT 5155
6094 Db GATCACCGCTCATGTACCCAGCTCTTCTTCTCGTGGCCAGCAGACCTATGT 6153
5156 QY GGTGCTCACTGTCATAAACCTCTTTATTGGCATCAATGGAAGCAGTGGCCACCTTGTCT 5215
6154 Db GGTGCTCACTGTCATAAACCTCTTTATTGGCATCAATGGAAGCAGTGGCCACCTTGTCT 6213
5216 QY TGAGCTCTTCTCTGATCAGAAAGCTGCGAGAGGCTGAGCCGAGTCTTGAACAGGTCTTCT 5275
6214 Db TGAGCTCTTCTCTGATCAGAAAGCTGCGAGAGGCTGAGCCGAGTCTTGAACAGGTCTTCT 6273
5276 QY TATCTTCCCCACCTTCTGCTTGGGCGGGGCTCATTTGACATGCTGGGACCCAGGCCAT 5335
6274 Db TATCTTCCCCACCTTCTGCTTGGGCGGGGCTCATTTGACATGCTGGGACCCAGGCCAT 6333
5336 QY GGCTGATGCTCTTGGAGGCTTGGGAGCAGGAGTTCAGTCCAGCTGCGCTGGAGGT 5395
6334 Db GGCTGATGCTCTTGGAGGCTTGGGAGCAGGAGTTCAGTCCAGCTGCGCTGGAGGT 6393
5396 QY GGTGGCAAGAACCTTGTGGCATGAGGAGTACAGGGGCCCCCTTCTTCTTCTTCACT 5455
6394 Db GGTGGCAAGAACCTTGTGGCATGAGGAGTACAGGGGCCCCCTTCTTCTTCTTCACT 6453
5456 QY ACTGCTCAGCACCGAAGCCTTCTGCGCACAGCCAGGCTGAGTCTTGGCACTCT 5515
6454 Db ACTGCTCAGCACCGAAGCCTTCTGCGCACAGCCAGGCTGAGTCTTGGCACTCT 6513
5516 QY GGGAGAGGAGCAGAGATGTAGCCCGTGAACGGGAGCGGTGGTCCAGAGCACCACA 5575
6514 Db GGGAGAGGAGCAGAGATGTAGCCCGTGAACGGGAGCGGTGGTCCAGAGCACCACA 6573
5576 QY GGGGATGTGTGTGTGAGGAACTTGAACCAAGGTATACCGTGGGCGAGAGATGCCAGC 5635
6574 Db GGGGATGTGTGTGTGAGGAACTTGAACCAAGGTATACCGTGGGCGAGAGATGCCAGC 6633
5636 QY TGTGACCGCTTGTGCTGGGATTTCCCTGTGTGTGTGTGTGGGCTGTGGGTGTA 5695
6634 Db TGTGACCGCTTGTGCTGGGATTTCCCTGTGTGTGTGTGTGGGCTGTGGGTGTA 6693
5696 QY TGGAGCAGGAGAGCTTCCAGCTTTCGATGCTGACCGGGGACACATTTGGGCGAGAGGG 5755
6694 Db TGGAGCAGGAGAGCTTCCAGCTTTCGATGCTGACCGGGGACACATTTGGGCGAGAGGG 6753
5756 QY CGAGGCTGTGTGCGAGGCCACAGCTGTGGCCCCGGAAACCCAGTGTGCGCACCTCAGCAT 5815

QY	3034	GTGGGATGCAACGGAAGCTGTCGTTGGCCATTGCTTTTGGGCGGCTCCCAAGTTGTTA	3093
Db	2941	GTGGGATGCAACGGAAGCTGTCGTTGGCCATTGCTTTTGGGCGGCTCCCAAGTTGTTA	3000
QY	3094	TCCTGGACGAGCCTACGGCTGGGGTGGATCTGCTTCCCGCCGGGTATTGGAGACTGC	3153
Db	3001	TCCTGGACGAGCCTACGGCTGGGGTGGATCTGCTTCCCGCCGGGTATTGGAGACTGC	3060
QY	3154	TGCTCAAAATACGAGAAAGTGGCAACGCTGATTCCTTCCACCCACACCTGGATGAGCGAG	3213
Db	3061	TGCTCAAAATACGAGAAAGTGGCAACGCTGATTCCTTCCACCCACACCTGGATGAGCGAG	3120
QY	3214	AGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGCTGTGGCTCCC	3273
Db	3121	AGCTGCTGGGAGACCGTGTGGCGTGTGGCAGGTGGCGCTGTGCTGCTGTGGCTCCC	3180
QY	3274	CACCTCTTCCTGCGCCGCTCACCTGGGGCTCCGGCTACTACCTGACGGCTGGTGAAGGCCCGCC	3333
Db	3181	CACCTCTTCCTGCGCCGCTCACCTGGGGCTCCGGCTACTACCTGACGGCTGGTGAAGGCCCGCC	3240
QY	3334	TGCCCTTGACCAACAATGAGAAAGCTGACACTGACATGAGAGGGAAGTGGGAACCAAGGC	3393
Db	3241	TGCCCTTGACCAACAATGAGAAAGCTGACACTGACATGAGAGGGAAGTGGGAACCAAGGC	3300
QY	3394	AGGAAAGAGAAATGGCAGCCAGGCGAGCAGAGTCCGCACTCCTCAGCTGCTGGCCCTGG	3453
Db	3301	AGGAAAGAGAAATGGCAGCCAGGCGAGCAGAGTCCGCACTCCTCAGCTGCTGGCCCTGG	3360
QY	3454	TACAGCACTGGGTGCCCGGGCCACGGCTGTGTGAGGAGCTGCCACACGAGCTGGTGTGG	3513
Db	3361	TACAGCACTGGGTGCCCGGGCCACGGCTGTGTGAGGAGCTGCCACACGAGCTGGTGTGG	3420
QY	3514	TGCTGCCCTACACGGGTGCCATGACGAGCAGCAGCTTCGCACTCTTCCCGAGAGCTAGACA	3573
Db	3421	TGCTGCCCTACACGGGTGCCATGACGAGCAGCAGCTTCGCACTCTTCCCGAGAGCTAGACA	3480
QY	3574	CGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGATCTCCGACACCAAGCCTCGAGAGA	3633
Db	3481	CGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGATCTCCGACACCAAGCCTCGAGAGA	3540
QY	3634	TCCTTCTGAAAGTGTGGAGAGTGTGCTCGGACACAGATATGAGAGTGGCAGCTGCG	3693
Db	3541	TCCTTCTGAAAGTGTGGAGAGTGTGCTCGGACACAGATATGAGAGTGGCAGCTGCG	3600
QY	3694	GGCAGCACTATGCAACGAGCATTTGCTGCCCTAGACGTTAACCTTCGGGCTCAAGATCGCG	3753
Db	3601	GGCAGCACTATGCAACGAGCATTTGCTGCCCTAGACGTTAACCTTCGGGCTCAAGATCGCG	3660
QY	3754	CACAGGAGCAGCGCTGGAGAACGGGAAACGAGCTGGGTACGCCCCAGAGACTGACCCAGG	3813
Db	3661	CACAGGAGCAGCGCTGGAGAACGGGAAACGAGCTGGGTACGCCCCAGAGACTGACCCAGG	3720
QY	3814	GCTCTGGGCGCAGACGCCGTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGCAGCTCC	3873
Db	3721	GCTCTGGGCGCAGACGCCGTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGCAGCTCC	3780
QY	3874	AGGCCCTGCTTCTCAAGCGCTTTCTGCTTGGCCCGCGCAGCCCGCGGCTGTTCGCC	3933
Db	3781	AGGCCCTGCTTCTCAAGCGCTTTCTGCTTGGCCCGCGCAGCCCGCGGCTGTTCGCC	3840
QY	3934	AGATCGTGTGCTGCCCTCTTTTGGGCGCTGGCCCTCGTGTTCAGGCTCATGTCGCTTC	3993
Db	3841	AGATCGTGTGCTGCCCTCTTTTGGGCGCTGGCCCTCGTGTTCAGGCTCATGTCGCTTC	3900
QY	3994	CTTTTCGGGCACTACCCGGCTCTCGGCTCAGTCCCACTACCTGATACGGTGTCTCAGGTGCT	4053
Db	3901	CTTTTCGGGCACTACCCGGCTCTCGGCTCAGTCCCACTACCTGATACGGTGTCTCAGGTGCT	3960
QY	4054	TCCTTCACTGAGGACGCCCAAGGGAACCTTGGACGTCGCCCGGCTGCTCGAGCGCTGCTGC	4113
Db	3961	TCCTTCACTGAGGACGCCCAAGGGAACCTTGGACGTCGCCCGGCTGCTCGAGCGCTGCTGC	4020

QY	4114	AGGAGCGAGACTGGAGGAGCCGCCAGTGCAGCATAGCTCCACAGGTTCTTCGGCACCCAG	4173
Db	4021	AGGAGCGAGACTGGAGGAGCCGCCAGTGCAGCATAGCTCCACAGGTTCTTCGGCACCCAG	4080
QY	4174	AAGTTCTCTGTAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGACCCACAGAGTCTCCAT	4233
Db	4081	AAGTTCTCTGTAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGACCCACAGAGTCTCCAT	4140
QY	4234	CCCCAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGCTGCCGAGTGCCTGGCTG	4293
Db	4141	CCCCAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGCTGCCGAGTGCCTGGCTG	4200
QY	4294	CAGCTGCTGGTCCCGCTCCCGCCAGGCACTGACCGGCTCTGGGGAGTGGTTTCAGAACC	4353
Db	4201	CAGCTGCTGGTCCCGCTCCCGCCAGGCACTGACCGGCTCTGGGGAGTGGTTTCAGAACC	4260
QY	4354	TGACAGGCGGGAACCTGTCTGACTTCTTGGTCAAGACCTACCCGCGCTTGGTGCAGG	4413
Db	4261	TGACAGGCGGGAACCTGTCTGACTTCTTGGTCAAGACCTACCCGCGCTTGGTGCAGG	4320
QY	4414	GCCTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACGAGGCTTCTCGCTGGGGGCC	4473
Db	4321	GCCTGAAGACTAAGAGTGGGTGAATGAGGTACAGGTACGAGGCTTCTCGCTGGGGGCC	4380
QY	4474	GAGACCCAGCCTGCCCTCGGGCCAAAGTGTGGGCGCTCAGTGGAGGAGTGTGTGGCGC	4533
Db	4381	GAGACCCAGCCTGCCCTCGGGCCAAAGTGTGGGCGCTCAGTGGAGGAGTGTGTGGCGC	4440
QY	4534	TGCTGAGTCCCTGCTGCGGGGGCCCTCGAACGCTGTCTGAAAAACCTCACAGCTGGG	4593
Db	4441	TGCTGAGTCCCTGCTGCGGGGGCCCTCGAACGCTGTCTGAAAAACCTCACAGCTGGG	4500
QY	4594	CTCAGAGCCTGAGTGTCTCAGGACAGTCTCAAGATCTGTTTCAACACAAAGGCTGCACT	4653
Db	4501	CTCAGAGCCTGAGTGTCTCAGGACAGTCTCAAGATCTGTTTCAACACAAAGGCTGCACT	4560
QY	4654	CCATGCTGGCTTTTGTCAACCGAGCAGCAACGCAATCTCTCGTGTCTCACCTCCGCCAG	4713
Db	4561	CCATGCTGGCTTTTGTCAACCGAGCAGCAACGCAATCTCTCGTGTCTCACCTCCGCCAG	4620
QY	4714	GCCCGCCCGCCACGCCCCACAGCATCACCACACTCAACACCCCTTGAACCTCACCAAGG	4773
Db	4621	GCCCGCCCGCCACGCCCCACAGCATCACCACACTCAACACCCCTTGAACCTCACCAAGG	4680
QY	4774	AGCAGCTGTCTGAGGCTGCACTGATGCTCTCTCGGTGAGACGCTCTCTCTCTCTCTG	4833
Db	4681	AGCAGCTGTCTGAGGCTGCACTGATGCTCTCTCGGTGAGACGCTCTCTCTCTCTCTG	4740
QY	4834	TGCTCTTTGGCCATGCTTTTGTCCCGCCAGCTTCACTCTTGTCTCTCATTTGAGGAGCGAG	4893
Db	4741	TGCTCTTTGGCCATGCTTTTGTCCCGCCAGCTTCACTCTTGTCTCTCATTTGAGGAGCGAG	4800
QY	4894	TCACCCGAGCCAAAGCACTGCACTCATGGGGGCTGTCCGCCACCCCTCTACTTGGCTG	4953
Db	4801	TCACCCGAGCCAAAGCACTGCACTCATGGGGGCTGTCCGCCACCCCTCTACTTGGCTG	4860
QY	4954	GCNACTTCTCTGGGACATGTGTAACTACTTGTGTGCGAGCATGCTGTGCTCTCATCT	5013
Db	4861	GCNACTTCTCTGGGACATGTGTAACTACTTGTGTGCGAGCATGCTGTGCTCTCATCT	4920
QY	5014	TTCTGGCCCTTCAGAGAGGCGCATATGTGGCCCTTCGCAACCTGCTGCTCTCTCTGCTGT	5073
Db	4921	TTCTGGCCCTTCAGAGAGGCGCATATGTGGCCCTTCGCAACCTGCTGCTCTCTCTGCTGT	4980
QY	5074	TGCTACTACTGTATGGCTGGTGCATCACACCGCTCATGTATGCCAGCCTCTCTCTCTCT	5133
Db	4981	TGCTACTACTGTATGGCTGGTGCATCACACCGCTCATGTATGCCAGCCTCTCTCTCTCT	5040
QY	5134	CGGTGCCAGCACAGCCTATGTGGTGTCTCACCTGCAATAACCTCTTTTATTTGGCATCAATG	5193
Db	5041	CGGTGCCAGCACAGCCTATGTGGTGTCTCACCTGCAATAACCTCTTTTATTTGGCATCAATG	5100
QY	5194	GAAGCATGGCCACCTTTTGTGCTTGTAGCTCTTCTCTGATCAGAAAGCTGCAGGAGGTGAGCC	5253

Db 5101 GAAGCATGCGCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGTGAGCC 5160
Qy 5254 GGATCTTGAACACAGCTCTTCTTATCTTCCCCCACTTCTGCTTGGCGCGGGGCTCATTTG 5313
Db 5161 GGATCTTGAACACAGCTCTTCTTATCTTCCCACTTCTGCTTGGCGCGGGGCTCATTTG 5220
Qy 5314 ACATGGTCGGGACACAGGCGCATGCTGATGCTCTTTGAGCGCTTGGAGACAGGCAAGTTCC 5373
Db 5221 ACATGGTCGGGACACAGGCGCATGCTGATGCTCTTTGAGCGCTTGGAGACAGGCAAGTTCC 5280
Qy 5374 AGTACCCCTCGCTGGGAGGTGTCGCAAGAACCTCTTGCCCATGCTGATACAGGGC 5433
Db 5281 AGTACCCCTCGCTGGGAGGTGTCGCAAGAACCTCTTGCCCATGCTGATACAGGGC 5340
Qy 5434 CCCTCTTCTCTTCTTCACTACTGCTGCAGCACCGAAGCAACTCTCTGCCACAGCCCA 5493
Db 5341 CCCTCTTCTCTTCTTCACTACTGCTGCAGCACCGAAGCAACTCTCTGCCACAGCCCA 5400
Qy 5494 GGGTAGGTCTCTGCCACTCTCTGGAGAGGAGGACGAGATGTAGCCCGTGAACGGGAGC 5553
Db 5401 GGGTAGGTCTCTGCCACTCTCTGGAGAGGAGGACGAGATGTAGCCCGTGAACGGGAGC 5460
Qy 5554 GGGTAGGTCTCTGCCACTCTCTGGAGAGGAGGACGAGATGTAGCCCGTGAACGGGAGC 5613
Db 5461 GGGTAGGTCTCTGCCACTCTCTGGAGAGGAGGACGAGATGTAGCCCGTGAACGGGAGC 5520
Qy 5614 ACCGTGGCAGAGGATGTCAGCTGTTGACCGCTTGTGCTGGGATTCCTCTGTGTAGT 5673
Db 5521 ACCGTGGCAGAGGATGTCAGCTGTTGACCGCTTGTGCTGGGATTCCTCTGTGTAGT 5580
Qy 5674 GTTTTGGGCTCTGCGTGTGAATGAGCAGGAGAACGCTCCACGTTTCGTCATGTGACGG 5733
Db 5581 GTTTTGGGCTCTGCGTGTGAATGAGCAGGAGAACGCTCCACGTTTCGTCATGTGACGG 5640
Qy 5734 GGGACATTTGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCGCACAGTGCGCCCGGGAAC 5793
Db 5641 GGGACATTTGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCGCACAGTGCGCCCGGGAAC 5700
Qy 5794 CCAGTGTGCGCACTCAGCATGGATGACTGCTTCAATCGATGCGCATCTTTGAGCTGC 5853
Db 5701 CCAGTGTGCGCACTCAGCATGGATGACTGCTTCAATCGATGCGCATCTTTGAGCTGC 5760
Qy 5854 TGACGGCGCGGAGCAGCTGGAGCTGCTTGCGCGCTGCGCGGTGTCGCGAGGCCCGAGG 5913
Db 5761 TGACGGCGCGGAGCAGCTGGAGCTGCTTGCGCGCTGCGCGGTGTCGCGAGGCCCGAGG 5820
Qy 5914 TTGCCACAGCGCTGGCTCAGCGCTTGGCGCTCTGGGACTCTCATGTGTACGACCGGC 5973
Db 5821 TTGCCACAGCGCTGGCTCAGCGCTTGGCGCTCTGGGACTCTCATGTGTACGACCGGC 5880
Qy 5974 CTGACGGCACCTACAGCGAGGGGAACAAACGCAAGCTGGCGACGCGCCCTGGCGCTGTTG 6033
Db 5881 CTGACGGCACCTACAGCGAGGGGAACAAACGCAAGCTGGCGACGCGCCCTGGCGCTGTTG 5940
Qy 6034 GGGACCCAGCGCTGCTGTTCTTGGACGAGCGGACCAACAGGATGGAATGGAATGGAATGGC 6093
Db 5941 GGGACCCAGCGCTGCTGTTCTTGGACGAGCGGACCAACAGGATGGAATGGAATGGAATGGC 6000
Qy 6094 GCTTCTTTTGAACAGCGCTTTTGGCGGTGTCGGGAGGCGGTTCAAGTGTCACTACCT 6153
Db 6001 GCTTCTTTTGAACAGCGCTTTTGGCGGTGTCGGGAGGCGGTTCAAGTGTCACTACCT 6060
Qy 6154 CCCATAGCATGGAGGTGTAAGCGCTCTGCTCGCGCTAGCCATCATGCTGATGATGGC 6213
Db 6061 CCCATAGCATGGAGGTGTAAGCGCTCTGCTCGCGCTAGCCATCATGCTGATGATGGC 6120
Qy 6214 GGTTCGCTGCTGGGAGCGCCGCAACATCTCAAGGGCAGATTCCGCGCGGTTCACACAC 6273
Db 6121 GGTTCGCTGCTGGGAGCGCCGCAACATCTCAAGGGCAGATTCCGCGCGGTTCACACAC 6180
Qy 6274 TGAACCTTCGCGGTGCCCGCGCAAGGTCCTCCAGCGGACGCGGCTTCTGTCGCGCGAGT 6333

Db 6181 TGAACCTTCGCGGTGCCCGCGCAAGGTCCTCCAGCGGACGCGCTTCTGTCGCGCGAGT 6240
Qy 6334 TCCCTGGGTCCGAGCTGCGCGGACATGAGGCGCGCTGCGCTTCCAGCTGCGCGCG 6393
Db 6241 TCCCTGGGTCCGAGCTGCGCGGACATGAGGCGCGCTGCGCTTCCAGCTGCGCGCG 6300
Qy 6394 GAGGCGCTGCGCGCTGCGCGCGCTTCTTTGAGAGCTGCGCGGTGCACGCGCAGACAG 6453
Db 6301 GAGGCGCTGCGCGCTGCGCGCGCTTCTTTGAGAGCTGCGCGGTGCACGCGCAGACAG 6360
Qy 6454 GCGTGGAGGACTTTCGCTGAGCCAGACGATGCTGGAGAGGATTTCTGTACTTCTCCA 6513
Db 6361 GCGTGGAGGACTTTCGCTGAGCCAGACGATGCTGGAGAGGATTTCTGTACTTCTCCA 6420
Qy 6514 AGGACGAGGGGAAGGACGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACC 6573
Db 6421 AGGACGAGGGGAAGGACGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACC 6480
Qy 6574 CCGCGCAGGCTGCAGCACCCCAAAACGCGTCAGCAGTTCTCTGATGACCCCTAGCCTG 6633
Db 6481 CCGCGCAGGCTGCAGCACCCCAAAACGCGTCAGCAGTTCTCTGATGACCCCTAGCCTG 6540
Qy 6634 CCGAGACTGCTCTGAGCCTTCCCTCCCTGCGGGGCGCGGGAGGCGCTTGGGAATGCG 6693
Db 6541 CCGAGACTGCTCTGAGCCTTCCCTCCCTGCGGGGCGCGGGAGGCGCTTGGGAATGCG 6600
Qy 6694 AAGGCGAAGTAGAGTGCCTAGGAGCCTGAGCTCAGCTCAGGAGGGGCTGTGCTCCT 6753
Db 6601 AAGGCGAAGTAGAGTGCCTAGGAGCCTGAGCTCAGCTCAGGAGGGGCTGTGCTCCT 6660
Qy 6754 GGAGAAATAAGAGAAAGGCTGGAGAGAGCCGTGCTGCTGTA 6797
Db 6661 GGAGAAATAAGAGAAAGGCTGGAGAGAGCCGTGCTGTA 6704

RESULT 7

US-10-775-920-3
; Sequence 3, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 6704
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-920-3

Query Match 98.2%; Score 6884.8; DB 17; Length 6704;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6692; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 94 GTGGGATAAGGAATGAGGTTTCAGAAAGGGGCGAGGAGTTGCCCGCAGCGCAGCCGACG 153
Db 1 GTGGGATAAGGAATGAGGTTTCAGAAAGGGGCGAGGAGTTGCCCGCAGCGCAGCCGACG 60
Qy 154 TCTTTCAGCCGACCGTTTGTCTGACCTCTCTGTCCCGTCCCGCCCGCAGTCTCACCATGG 213
Db 61 TCTTTCAGCCGACCGTTTGTCTGACCTCTCTGTCCCGTCCCGTCCCGCAGTCTCACCATGG 120
Qy 214 CTTTCTGGACACAGCTGATGCTGCTGTGGAAGAATTTTCATGATATCGCCGAGACAGC 273
Db 121 CTTTCTGGACACAGCTGATGCTGCTGTGGAAGAATTTTCATGATATCGCCGAGACAGC 180

Qy	274	CGGTCCAGACTCCTGGTGGAAATTGCTGTGGCCCTCTCTTCCTCTTCTTCAATCTGTGTGCTG	333
Db	181	CGGTCCAGACTCCTGGTGGAAATTGCTGTGGCCCTCTCTTCCTCTTCTTCAATCTGTGTGCTG	240
Qy	334	TTTCGCCACTCCACCCGCCCTGGAGCACCATGAAATGCCACTTCCCAAAACAAGCCACTGC	393
Db	241	TTTCGCCACTCCACCCGCCCTGGAGCACCAATGAATGCCTTTCCTCCAAAACAAGCCACTGC	300
Qy	394	CATCGGGGGCACCCTGTGCCCTCGAGGCTCCAGAGGTCTCATCTGTAATGTGAAACAACACTGCT	453
Db	301	CATCGGGGGCACCCTGTGCCCTCGAGGCTCATCTGTAATGTGAAACAACACTGCT	360
Qy	454	TTTCGCAGCTGACACCGGGCGAGAGGCCCGGGCGCTGTAGACAACTTCAACGACTCCTCTGG	513
Db	361	TTTCGCAGCTGACACCGGGCGAGAGGCCCGGGCGCTGTAGCAACTTCAACGACTCCTCTGG	420
Qy	514	TCTTCCGGCTGTAGCCGATGCCCGCACTGTGCTGGGAGGGGCCAGTGCCTCCACAGGACGC	573
Db	421	TCTTCCGGCTGTAGCCGATGCCCGCACTGTGCTGGGAGGGGCCAGTGCCTCCACAGGACGC	480
Qy	574	TGGCTTGGCTTAGGGAAGCTGATGCCACGCTAGGGCTTGCAACGAGACACGGCCCCAGCCTC	633
Db	481	TGGCTTGGCTTAGGGAAGCTGATGCCACGCTAGGGCTTGCAACGAGACACGGCCCCAGCCTC	540
Qy	634	AACCAACCAAGCAGTCTCCACTGAGAACCAACCCATGCTGGATGTGCGGGAGCTGTGACGT	693
Db	541	AACCAACCAAGCAGTCTCCACTGAGAACCAACCCATGCTGGATGTGCGGGAGCTGTGACGT	600
Qy	694	CACTGCTTGGCAACCGAATCCCTTGGGGTTGGCACTGGGGCAAGCCACGAGAGCCCTTGCACA	753
Db	601	CACTGCTTGGCAACCGAATCCCTTGGGGTTGGCACTGGGGCAAGCCACGAGAGCCCTTGCACA	660
Qy	754	GCTTGTGGAGGCGCTGAGGACCTGSCCCAGGAGCTCTTGGCGCTGCGAGACCTGTGTGG	813
Db	661	GCTTGTGGAGGCGCTGAGGACCTTGGCCACGAGACTCTTGGCGCTGCGAGACCTGTGTGG	720
Qy	814	AGCTTCGGGCACCTGTCGACAGACCCCGAGGACACAGCGGCCCTCGAGTGTGTGTGAC	873
Db	721	AGCTTCGGGCACCTGTCGACAGACCCCGAGGACACAGCGGCCCTCGAGTGTGTGTGAC	780
Qy	874	AGGCCCTGTGAGTGTGACGGGAACCTAGACGACAAGTGGGGCCCTCTCCCTCAACTGTGTACG	933
Db	781	AGGCCCTGTGAGTGTGACGGGAACCTAGACGACAAGTGGGGCCCTCTCCCTCAACTGTGTACG	840
Qy	934	AGGCTAGTGACCTGATGAGCTGTGTGGGACAGGACCAAGATCGCCCTGCCACACAGACA	993
Db	841	AGGCTAGTGACCTGATGAGCTGTGTGGGACAGGACCAAGATCGCCCTGCCACACAGACA	900
Qy	994	GCCTGAGCCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCCGCTGTCCGCC	1053
Db	901	GCCTGAGCCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCCGCTGTCCGCC	960
Qy	1054	TGCTCTGGAGACGCTGAGCCTGTGATCTCTCGGGAAAGCTACTCTTTGGACACAGATACAC	1113
Db	961	TGCTCTGGAGACGCTGAGCCTGTGATCTCTCGGGAAAGCTACTCTTTGGACACAGATACAC	1020
Qy	1114	CTTTTATCCCGAAGCTCATGSCCAGGTGACCCGACCTTTCGAGGACTCACCTGTCTGA	1173
Db	1021	CTTTTATCCCGAAGCTCATGSCCAGGTCAACCCGACCTTTCGAGGAGCTCACCTGTCTGA	1080
Qy	1174	GGGATGTCCGGAGGTGTGGAGATGTGTGGACCCCGGATCTTTCACCTTTCATGAACGACA	1233
Db	1081	GGGATGTCCGGAGGTGTGGAGATGTGTGGACCCCGGATCTTTCACCTTTCATGAACGACA	1140
Qy	1234	GTTTCCAATGTGGCCATGCTGACGGGCTCTGACAGATGACAGGATGAAGGAAGAGGACG	1293
Db	1141	GTTTCCAATGTGGCCATGCTGACGGGCTCTGACAGATGACAGGATGAAGGAAGAGGACG	1200
Qy	1294	CCAGACTTGGAGGCGGGACACCATGGAGGCCCTCGCATCTTTTCTGGAACCTTGGAGCG	1353
Db	1201	CCAGACTTGGAGGCGGGACACCATGGAGGCCCTCGCATCTTTTCTGGAACCTTGGAGCG	1260
Qy	1354	GTGGCTACAGCTGGCAGGACGCACACGCTGATGTGGGGCACCTTGGTGGGACACGCTGGGCC	1413

	Db																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
--	----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 2341 CGCTCTACGGGCTCCGCCACTGGTACTGGGAAGCTGTGTGCCAGGCGCAGTACGGGATCC 2400
Qy 2494 CTGAACCATGGAATTTCTTTTCGGAGGAGCTACTGTGTGGGACCTCGGGCCCCCAAGA 2553
Db 2401 CTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGTGTGGGACCTCGGGCCCCCAAGA 2460
Qy 2554 GTCCAGCCCCCTTGCCCAACCCCGCTGGACCCAAAGGTGCTGTGAAGGACCGGCCG 2613
Db 2461 GTCCAGCCCCCTTGCCCAACCCCGCTGGACCCAAAGGTGCTGTGAAGGACCGGCCG 2520
Qy 2614 GCCTGAGTCTTGGCGTATCCGTTTCGACCTGGAGAAGCGCTTCTCTGGAAGCCGCGACG 2673
Db 2521 GCCTGAGTCTTGGCGTCTCCGTTTCGACCTGGAGAAGCGCTTCTCTGGAAGCCGCGACG 2580
Qy 2674 CAGCCCTCGGGGGCTCAGCCCTGGAATTTCTACAGGGCCACATCACCCCTTCTCTGGGCC 2733
Db 2581 CAGCCCTCGGGGGCTCAGCCCTGGAATTTCTACAGGGCCACATCACCCCTTCTCTGGGCC 2640
Qy 2734 ACAACGGGGCCGGAAGACCAACCCCTGTCTGATCTTGAAGTGGCTCTTCCACCCAGTG 2793
Db 2641 ACAACGGGGCCGGAAGACCAACCCCTGTCTGATCTTGAAGTGGCTCTTCCACCCAGTG 2700
Qy 2794 GTGGCTCTGCTTCACTCATCTGGGCCACAGACGTCCTCCGCTCCAGATGSCGCGCATCCGGCCCC 2853
Db 2701 GTGGCTCTGCTTCACTCATCTGGGCCACAGACGTCCTCCGCTCCAGATGSCGCGCATCCGGCCCC 2760
Qy 2854 ACCTGGGGCTCTGCTCCTCAGTACAACTGCTGTTTGAATGCTGACCTGACCGTGGACGAGACG 2913
Db 2761 ACCTGGGGCTCTGCTCCTCAGTACAACTGCTGTTTGAATGCTGACCTGACCGTGGACGAGCAG 2820
Qy 2914 TCTGTTTCTATGGGGCGCTGAAGGTCGTGAGTGGCGCTGTGAGTGGCCCCCGAGCAGGACC 2973
Db 2821 TCTGTTTCTATGGGGCGCTGAAGGTCGTGAGTGGCGCTGTGAGTGGCCCCCGAGCAGGACC 2880
Qy 2974 GTCTGCTGCAGATGTGGGGCTGTCTCCAGCAGAGTGTCCAGTCTGCGACCTCTCTG 3033
Db 2881 GTCTGCTGCAGATGTGGGGCTGTCTCCAGCAGAGTGTCCAGTCTGCGACCTCTCTG 2940
Qy 3034 GTGGGATGCAACGGGAAGCTGTCCTGGGCCATTTGCTTTGGGGGGCTCCCAAGTTGTTA 3093
Db 2941 GTGGGATGCAACGGGAAGCTGTCCTGGGCCATTTGCTTTGGGGGGCTCCCAAGTTGTTA 3000
Qy 3094 TCCTGGACAGCTACCGCTCGGCTGAGATCTGCTTCCCGCGCGGATTTGGAGTGC 3153
Db 3001 TCCTGGACAGCTACCGCTCGGCTGAGATCTGCTTCCCGCGCGGATTTGGAGTGC 3060
Qy 3154 TGCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCACTGATGAGGACG 3213
Db 3061 TGCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCACTGATGAGGACG 3120
Qy 3214 AGCTGCTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGCTGTGGCTCCC 3273
Db 3121 AGCTGCTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGCTGTGGCTCCC 3180
Qy 3274 CACTTCTCTCGCGCTCACTTGGCTCCGCTACTACTGACGCTGTGTAAGGCCCGGCC 3333
Db 3181 CACTTCTCTCGCGCTCACTTGGCTCCGCTACTACTGACGCTGTGTAAGGCCCGGCC 3240
Qy 3334 TGCCCTTGACCAACCAATGAGAGGCTGACATGATGAGGGCAGTGTGACACCCAGGC 3393
Db 3241 TGCCCTTGACCAACCAATGAGAGGCTGACATGATGAGGGCAGTGTGACACCCAGGC 3300
Qy 3394 AGGAAAGAAATGGGACCGAGGACGAGTGGGACCTCTCCTCAGCTGCTGGCCCTGG 3453
Db 3301 AGGAAAGAAATGGGACCGAGGACGAGTGGGACCTCTCCTCAGCTGCTGGCCCTGG 3360
Qy 3454 TACAGCACTGGGTGCCCGGGCAACGCTGGTGGAGGAGCTCCCAACAGAGCTGGTGTGG 3513
Db 3361 TACAGCACTGGGTGCCCGGGCAACGCTGGTGGAGGAGCTCCCAACAGAGCTGGTGTGG 3420
Qy 3514 TGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCTCTTCGAGAGCTAGACA 3573
Db 3421 TGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCTCTTCGAGAGCTAGACA 3480

Qy 3574 CGCGCTCGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCGAGCTTCGAGGAGA 3633
Db 3481 CGCGCTCGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCGAGCTTCGAGGAGA 3540
Qy 3634 TCTTCTCTGAAGGTGGTGGAGAGTGTCTGTGGGACACAGATATGGAGATGGCAGTGGC 3693
Db 3541 TCTTCTCTGAAGGTGGTGGAGAGTGTCTGTGGGACACAGATATGGAGATGGCAGTGGC 3600
Qy 3694 GGACAGCACTATGCACAGGCAATGCTGGCTTAGAGCTAACTCTGGGCTCAAGATGGCCG 3753
Db 3601 GGACAGCACTATGCACAGGCAATGCTGGCTTAGAGCTAACTCTGGGCTCAAGATGGCCG 3660
Qy 3754 CACAGGACAGCGCTGGAGAACCGGGAAACAGCTGGGTTCAGCCCCCAGAGACTGACCAGG 3813
Db 3661 CACAGGACAGCGCTGGAGAACCGGGAAACAGCTGGGTTCAGCCCCCAGAGACTGACCAGG 3720
Qy 3814 GCTCTGGGCCACAGCGCGTGGGCCCGGGTACAGGGCTGGGGCACTGACCCGCCAGAGCTCC 3873
Db 3721 GCTCTGGGCCACAGCGCGTGGGCCCGGGTACAGGGCTGGGGCACTGACCCGCCAGAGCTCC 3780
Qy 3874 AGGCCCTGCTTCTCAAGCGCTTCTGTGCTTGCCCGCGCAGCGCGCGCTGTTCGGCC 3933
Db 3781 AGGCCCTGCTTCTCAAGCGCTTCTGTGCTTGCCCGCGCAGCGCGCGCTGTTCGGCC 3840
Qy 3934 AGATCGTCTGCTGCTCCCTCTTTGTGGGCCCTGGGCCCTCGTGTTCAGCCCTCATTCGTCCTC 3993
Db 3841 AGATCGTCTGCTGCTCCCTCTTTGTGGGCCCTGGGCCCTCGTGTTCAGCCCTCATTCGTCCTC 3900
Qy 3994 CTTTTCGGGCACTACCGGCTCTCGGGCTCAGTCCCAACCATGTACGGTCTCAGGTGCTCT 4053
Db 3901 CTTTTCGGGCACTACCGGCTCTCGGGCTCAGTCCCAACCATGTACGGTCTCAGGTGCTCT 3960
Qy 4054 TCTTCAGTGAAGACGCCCCAGGGGACCTGTGACGCTGCCCGCTCTCGAGGGCGCTGCTGC 4113
Db 3961 TCTTCAGTGAAGACGCCCCAGGGGACCTGTGACGCTGCCCGCTCTCGAGGGCGCTGCTGC 4020
Qy 4114 AGGAGGACGAGCTGGAGGAGCCCCAGTGCAGCATAGTCTCCACACAGGTTCTCGGCACCCAG 4173
Db 4021 AGGAGGACGAGCTGGAGGAGCCCCAGTGCAGCATAGTCTCCACACAGGTTCTCGGCACCCAG 4080
Qy 4174 AAGTCTCTGTGAAGTGGCCAAAGTCTTTGGCCAGTGGCACTGAGACCCACAGTCTCCAT 4233
Db 4081 AAGTCTCTGTGAAGTGGCCAAAGTCTTTGGCCAGTGGCACTGAGACCCACAGTCTCCAT 4140
Qy 4234 CCCAGCTGCTGAGTGTAGCCAGCGCGCTGCCCGCTCTGCTGCCAGCTGCTGCCGGCTG 4293
Db 4141 CCCAGCTGCTGAGTGTAGCCAGCGCGCTGCCCGCTCTGCTGCCAGCTGCTGCCGGCTG 4200
Qy 4294 CAGCTGTGTCTCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGGTTTCAGAAACC 4353
Db 4201 CAGCTGTGTCTCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGGTTTCAGAAACC 4260
Qy 4354 TGAACGCGGAAACCTGTCTGACTTCTGTGTGAACCTACCCGCGCTGTGTGGCCAGG 4413
Db 4261 TGAACGCGGAAACCTGTCTGACTTCTGTGTGAACCTACCCGCGCTGTGTGGCCAGG 4320
Qy 4414 GCCTGAAGACTTAAGAAGTGGTGAATGAGGTGAGTACGGAGGCTTCTGCTGGGGGCC 4473
Db 4321 GCCTGAAGACTTAAGAAGTGGTGAATGAGGTGAGTACGGAGGCTTCTGCTGGGGGCC 4380
Qy 4474 GAGACCCAGGCTCTCCCTCCCGCCAAAGTGTGGGCCCTCAGTGGAGGAGTGTGGGGGCC 4533
Db 4381 GAGACCCAGGCTCTCCCTCCCGCCAAAGTGTGGGCCCTCAGTGGAGGAGTGTGGGGGCC 4440
Qy 4534 TGCTGAGTCTCCCTCGCTGGGGGCCCTTCGACCGTGTCTGAAACCTCAAGCCTGGG 4593
Db 4441 TGCTGAGTCTCCCTCGCTGGGGGCCCTTCGACCGTGTCTGAAACCTCAAGCCTGGG 4500
Qy 4594 CTCAAGCCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAAGGCTGGCACT 4653
Db 4501 CTCAAGCCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAAGGCTGGCACT 4560

QY	4654	CCATGGTGGCTTTTGTCAACCGAGCCAGCAACGCAATCTCTCGTGTCTACCTGCCCCAG	4713
Db	4561	CCATGGTGGCTTTTGTCAACCGAGCCAGCAACGCAATCTCTCGTGTCTACCTGCCCCAG	4620
QY	4714	GCCCGGCCCGCCACGCGCCACAGCATCAACACTCAACACCCCTTGAACCTCACCAAGG	4773
Db	4621	GCCCGGCCCGCCACGCGCCACAGCATCAACACTCAACACCCCTTGAACCTCACCAAGG	4680
QY	4774	AGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGAGCTCTCTGCTCATCTGTG	4833
Db	4681	AGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGAGCTCTCTGCTCATCTGTG	4740
QY	4834	TGCTCTTTGGCATGTCTCTTTGTCCTCGGCCAGCTTCACTCTTGTCTCAATTGAGAGCGAG	4893
Db	4741	TGCTCTTTGGCATGTCTCTTTGTCCTCGGCCAGCTTCACTCTTGTCTCAATTGAGAGCGAG	4800
QY	4894	TCACCGGAGCAAGCAACCTGCAGCTCATGTGGGGGCTGTCCCCCACCCTCTACTGGCTTG	4953
Db	4801	TCACCGGAGCAAGCAACCTGCAGCTCATGTGGGGGCTGTCCCCCACCCTCTACTGGCTTG	4860
QY	4954	GCAACTTTCTCTGGGACATGTGTAACTAATGTGTGGCAGCATGTGTGCTCATCT	5013
Db	4861	GCAACTTTCTCTGGGACATGTGTAACTAATGTGTGGCAGCATGTGTGCTCATCT	4920
QY	5014	TTCTGGCTTTCAGACGAGGGCATATGTGGCCCTGCGCAACCTGCTCTCTCTCTCT	5073
Db	4921	TTCTGGCTTTCAGACGAGGGCATATGTGGCCCTGCGCAACCTGCTCTCTCTCTCT	4980
QY	5074	TGCTACTTGTATGTGTGTGTCATCACACCGCTCATGTACCCAGCCTCTCTCTCTCT	5133
Db	4981	TGCTACTTGTATGTGTGTGTCATCACACCGCTCATGTACCCAGCCTCTCTCTCTCT	5040
QY	5134	CCGTGCCACGACAGCCTATGTGGTGTCACTGTGATTAACCTCTTTATTTGGCATCAATG	5193
Db	5041	CCGTGCCACGACAGCCTATGTGGTGTCACTGTGATTAACCTCTTTATTTGGCATCAATG	5100
QY	5194	GAAGCATGGCCACTTTGTGCTTGTAGCTTCTCTGATCAGAAGCTGCAGAGGTGAGCC	5253
Db	5101	GAAGCATGGCCACTTTGTGCTTGTAGCTTCTCTGATCAGAAGCTGCAGAGGTGAGCC	5160
QY	5254	GGATCTTTGAAACAGGCTTCTCTTATCTTCCCCCACTTCTGCTTGGGCGGGGCTCATTTG	5313
Db	5161	GGATCTTTGAAACAGGCTTCTCTTATCTTCCCCCACTTCTGCTTGGGCGGGGCTTATTG	5220
QY	5314	ACATGTGCGGAAACGAGGCCATGGCTGTATGCTTTGTAGCGCTTGGAGACAGCAGTTCC	5373
Db	5221	ACATGTGCGGAAACGAGGCCATGGCTGTATGCTTTGTAGCGCTTGGAGACAGCAGTTCC	5280
QY	5374	AGTCACCCCTGCGCTGGAGGTGCTCGGCAAGAACCTCTTTGGCCATGGTGATACAGGGG	5433
Db	5281	AGTCACCCCTGCGCTGGAGGTGCTCGGCAAGAACCTCTTTGGCCATGGTGATACAGGGG	5340
QY	5434	CCCTCTTCTCTTCTTCACTACTGTGTGAGCAGCAGCAAGCACTCTCTGCCACAGCCCA	5493
Db	5341	CCCTCTTCTCTTCTTCACTACTGTGTGAGCAGCAGCAAGCACTCTCTGCCACAGCCCA	5400
QY	5494	GGGTGAGGTCTGTGCCACTCTCTGGGAGAGGAGCAGAGATGTAGCCCGTGAACGGGAGC	5553
Db	5401	GGGTGAGGTCTGTGCCACTCTCTGGGAGAGGAGCAGAGATGTAGCCCGTGAACGGGAGC	5460
QY	5554	GGGTGCTTCAAGGAGCAGCCAGGGGATGTGTGTGTGTGTGAGCACTTGAACCAAGGTAT	5613
Db	5461	GGGTGCTTCAAGGAGCAGCCAGGGGATGTGTGTGTGTGTGAGCACTTGAACCAAGGTAT	5520
QY	5614	ACCGTGGGACAGAGATGCCAGCTGTGTGACCGCTTGTGCTTGGGGAATTCCTCCCTGTGAGT	5673
Db	5521	ACCGTGGGACAGAGATGCCAGCTGTGTGACCGCTTGTGCTTGGGGAATTCCTCCCTGTGAGT	5580
QY	5674	GTTTTGGGCTGTCTGGGTGTGAATGGAGCAGGGAAGAGCTCCAGTTTTCGATGGTACGG	5733
Db	5581	GTTTTGGGCTGTCTGGGTGTGAATGGAGCAGGGAAGAGCTCCAGTTTTCGATGGTACGG	5640
QY	5734	GGGACACATTTGGCCAGCAGGGGCGAGGCTGTGTGGCAGGCCACAGCGTGGCCGGGAAC	5793

Db	5641	GGGACACATTTGGCCAGCAGGGCGAGGCTGTGTGGCAGGCCACACGCTGGCCCGGGAAC	5700
QY	5794	CCAGTCTCTGGCACACTCAGCATGGGATACTGCCCTCAATCCGATGCATCTTTGAGCTGC	5853
Db	5701	CCAGTCTCTGGCACACTCAGCATGGGATACTGCCCTCAATCCGATGCATCTTTGAGCTGC	5760
QY	5854	TGACGGGCGCGAGACACCTGGAGCTGTGTGGCGCTGCGCGCTGCGGTGTCGCGAGGCCAGG	5913
Db	5761	TGACGGGCGCGAGACACCTGGAGCTGTGTGGCGCTGCGCGCTGCGGTGTCGCGAGGCCAGG	5820
QY	5914	TTGCCCCAGACCCGTGTGCTCAGGCTTGGCGCTCTGCGAATCTCATGTGTACGACAGACGGC	5973
Db	5821	TTGCCCCAGACCCGTGTGCTCAGGCTTGGCGCTCTGCGAATCTCATGTGTACGACAGACGGC	5880
QY	5974	CTGACGGCACTTACAGCGGAGGAAACAAACGCAAGCTGTGCGACGCGCTCTGGCGCTGGTTG	6033
Db	5881	CTGACGGCACTTACAGCGGAGGAAACAAACGCAAGCTGTGCGACGCGCTCTGGCGCTGGTTG	5940
QY	6034	GGGACCCAGCCGTGTGTCTTCTGGACGAGCGCCACACAGGCATGGACCCAGCGCGGGC	6093
Db	5941	GGGACCCAGCCGTGTGTCTTCTGGACGAGCGCCACACAGGCATGGACCCAGCGCGGGC	6000
QY	6094	GCTTCTCTTTGGAACAGCCTTTTGGCGCTGTGTGCGGAGGGCCGTTTCACTGATGCTCACCT	6153
Db	6001	GCTTCTCTTTGGAACAGCCTTTTGGCGCTGTGTGCGGAGGGCCGTTTCACTGATGCTCACCT	6060
QY	6154	CCCATAGCATGTGAGGAGTGTGAAGCGCTCTGTCTGCGCCTAGCCATCATGTGAATGGGC	6213
Db	6061	CCCATAGCATGTGAGGAGTGTGAAGCGCTCTGTCTGCGCCTAGCCATCATGTGAATGGGC	6120
QY	6214	GGTTCCGCTGCTGGGACGCGCCCAACATCTCAAGGGCAGATTCGCGGGGGTGCACAC	6273
Db	6121	GGTTCCGCTGCTGGGACGCGCCCAACATCTCAAGGGCAGATTCGCGGGGGTGCACAC	6180
QY	6274	TGACCCCTGCGGGTGCCCGCGCAAGGTCCAGCGCGCAGCGGCTTTCTGTGCGCGCGAGT	6333
Db	6181	TGACCCCTGCGGGTGCCCGCGCAAGGTCCAGCGCGCAGCGGCTTTCTGTGCGCGCGAGT	6240
QY	6334	TCCCTGGGTGAGAGCTGTGCGAGGCAATGTGAGGCGCGCTTCTGAGCTGCGCGCGG	6393
Db	6241	TCCCTGGGTGAGAGCTGTGCGAGGCAATGTGAGGCGCGCTTCTGAGCTGCGCGCGG	6300
QY	6394	GAGGGCGCTGCGCCCTTGGCGCGCTTGGAGAGCTGTGCGGTGCGCGCGAGAGCAGC	6453
Db	6301	GAGGGCGCTGCGCCCTTGGCGCGCTTGGAGAGCTGTGCGGTGCGCGCGAGAGCAGC	6360
QY	6454	GCGTGAGGACCTTTTCCGTGAGCCAGACGATGTGAGGAGGTATTTCTTGTACTTCTCCA	6513
Db	6361	GCGTGAGGACCTTTTCCGTGAGCCAGACGATGTGAGGAGGTATTTCTTGTACTTCTCCA	6420
QY	6514	AGGACCAAGGGGAGGACGAGGACACCGAAGAGCAGAGGAGGAGGTGGAGTGGAACC	6573
Db	6421	AGGACCAAGGGGAGGACGAGGACACCGAAGAGCAGAGGAGGAGGTGGAGTGGAACC	6480
QY	6574	CCGCGCAGGCTGTGAGCAGCCCAACCGCTGAGCAGTTCCTCGATGACCTAGCAGCTG	6633
Db	6481	CCGCGCAGGCTGTGAGCAGCCCAACCGCTGAGCAGTTCCTCGATGACCTAGCAGCTG	6540
QY	6634	CCGAGACTGTGCTCTGAGCCTTCCCTTCCCTGCGGGGCGCGGGGAGGCCCTGGGAATGGC	6693
Db	6541	CCGAGACTGTGCTCTGAGCCTTCCCTTCCCTGCGGGGCGCGGGGAGGCCCTGGGAATGGC	6600
QY	6694	AAGGGCAAGGTAGAGTGTCTAGAGCCCTGGAATCTCAGGCTGGCAGAGGGGCTGGTGCCT	6753
Db	6601	AAGGGCAAGGTAGAGTGTCTAGAGCCCTGGAATCTCAGGCTGGCAGAGGGGCTGGTGCCT	6660
QY	6754	GGGAAAAATAAGAGAGGCTGGAGAGAGCCGTGCTGGTGAAA	6797
Db	6661	GGGAAAAATAAGAGAGGCTGGAGAGAGCCGTGCTGGTGAAA	6704

US-10-775-920-4
; Sequence 4, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Meigen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4

; LENGTH: 6588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-4
Query Match 96.5%; Score 6565.6; DB 17; Length 6588;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6574; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy	210	ATGGCCCTTCTGGACACAGCTGATGCTGCTCTGCTCTGGAAGAAATTTTCATGTATCGCCGGAGA	269
Db	1	ATGGCCCTTCTGGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCCGGAGA	60
Qy	270	CAGCGGTCCAGACTCTGGTGAATGTGTGGCCCTCTCTTCCTTCTTATCTGTGTG	329
Db	61	CAGCGGTCCAGACTCTGGTGAATGTGTGGCCCTCTCTTCCTTCTTATCTGTGTG	120
Qy	330	GCTGTTCGCCACTCCACCGCCCTGGAGACACATGAATGCCATCTCCCAACAAAGCCA	389
Db	121	GCTGTTCGCCACTCCACCGCCCTGGAGACACATGAATGCCATCTCCCAACAAAGCCA	180
Qy	390	CTGCCATCGGGGGCACCGTCCCTGGCTCAGGGTCTCATCTGTAATGTGAACAAACACC	449
Db	181	CTGCCATCGGGGGCACCGTCCCTGGCTCAGGGTCTCATCTGTAATGTGAACAAACACC	240
Qy	450	TGCTTTCCGAGCTGACACCGGGGAGAGAGCCCGGGCCCTGAGCAATTCACAGACTCC	509
Db	241	TGCTTTCCGAGCTGACACCGGGGAGAGAGCCCGGGCCCTGAGCAATTCACAGACTCC	300
Qy	510	CTGCTCTCCCGCTAGCGATGCCCGCACTGTGCTGGAGGGGGCCAGTGCACACAGG	569
Db	301	CTGCTCTCCCGCTAGCGATGCCCGCACTGTGCTGGAGGGGGCCAGTGCACACAGG	360
Qy	570	ACGCTGGCTGGCTAGGGAAGCTGATCGCCACGCTGAGGGTGCACGACGACCGGCCAG	629
Db	361	ACGCTGGCTGGCTAGGGAAGCTGATCGCCACGCTGAGGGTGCACGACGACCGGCCAG	420
Qy	630	CCTCAACCAACCAAGACTCTCCACTGAAACACCCATGCTGGATGTCCGGAGTGTCTG	689
Db	421	CCTCAACCAACCAAGACTCTCCACTGAAACACCCATGCTGGATGTCCGGAGTGTCTG	480
Qy	690	ACGTCACCTGCGCAGCGAATCCCTGGGGTGGCACTGGGCCAAGCCAGAGGCCCTTG	749
Db	481	ACGTCACCTGCGCAGCGAATCCCTGGGGTGGCACTGGGCCAAGCCAGAGGCCCTTG	540
Qy	750	CACAGCTGTTGGAGGCCGCTGAGGACCTGGCCACAGAGCTCCTGGGCGCTGCGAGCCTG	809
Db	541	CACAGCTGTTGGAGGCCGCTGAGGACCTGGCCACAGAGCTCCTGGGCGCTGCGAGCCTG	600
Qy	810	GTGAGCTTCGGGCACTGCTCAGAGACCCCGAGGGACACGCGGCCCTCTGGAGTTGCTG	869
Db	601	GTGAGCTTCGGGCACTGCTCAGAGACCCCGAGGGACACGCGGCCCTCTGGAGTTGCTG	660
Qy	870	TCAGAGGCCCTCTGCAAGTGTACGGGACCTAGCAGCACAGTGGGCCCTCCTCACTGG	929
Db	661	TCAGAGGCCCTCTGCAAGTGTACGGGACCTAGCAGCACAGTGGGCCCTCCTCACTGG	720

Qy	930	TACGAGGCTAGTACCTGATCGAGCTGTGTGGGCGAGGACCCAGAAATCCGCTTCGCAGAC	989
Db	721	TACGAGGCTAGTACCTGATCGAGCTGTGTGGGCGAGGACCCAGAAATCCGCTTCGCAGAC	780
Qy	990	AGCAGCTGAGCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCGCTGTCC	1049
Db	781	AGCAGCTGAGCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCACCCGCTGTCC	840
Qy	1050	CGCTGCTCTCGAGACGCGCTGTAAGCCTCTGATCTCTCGGGAAGCTACTCTTTGCACAGAT	1109
Db	841	CGCTGCTCTCGAGACGCGCTGTAAGCCTCTGATCTCTCGGGAAGCTACTCTTTGCACAGAT	900
Qy	1110	ACACTTTTACCCGGAAGCTCATGCGCCAGGTGAACCGGACCTTCGAGAGCTCACTCTG	1169
Db	901	ACACTTTTACCCGGAAGCTCATGCGCCAGGTCAACCGGACCTTCGAGAGCTCACTCTG	960
Qy	1170	CTGAGGATGTCCGGAGGTGTGGAGATGCTGGGACCCCGGATCTTTCACCTTCAATGAC	1229
Db	961	CTGAGGATGTCCGGAGGTGTGGAGATGCTGGGACCCCGGATCTTTCACCTTCAATGAC	1020
Qy	1230	GACAGTTCCAAATGTGGCCATGCTGAGCGGCTCTCTGAGATGTCAGGATGAAGGAAGAGG	1289
Db	1021	GACAGTTCCAAATGTGGCCATGCTGAGCGGCTCTCTGAGATGTCAGGATGAAGGAAGAGG	1080
Qy	1290	CAGCCACAGACCTGGAGCCGGGACACATGAGGCGCTTCGATCTTTCTGACACCTGGG	1349
Db	1081	CAGCCACAGACCTGGAGCCGGGACACATGAGGCGCTTCGATCTTTCTGACACCTGGG	1140
Qy	1350	AGCGTGTGCTACAGCTGGCAGAGCGCACACCTGATGTGGGGCACCTGGTGGGACGCTG	1409
Db	1141	AGCGTGTGCTACAGCTGGCAGAGCGCACACCTGATGTGGGGCACCTGGTGGGACGCTG	1200
Qy	1410	GGCCAGTGAAGGAGTGCCTGCTTGGACAAGCTTGAAGGGGCGCACCTCAGAGGACGCC	1469
Db	1201	GGCCAGTGAAGGAGTGCCTGCTTGGACAAGCTTGAAGGGGCGCACCTCAGAGGACGCC	1260
Qy	1470	CTGTGTGCGGGGCCCTGCAACTGCTCGCGAAACATCGATTTCTGGGCGGGTCTGCTTC	1529
Db	1261	CTGTGTGCGGGGCCCTGCAACTGCTCGCGAAACATCGATTTCTGGGCGGGTCTGCTTC	1320
Qy	1530	TTGGGACCTGAGGACTCTTTCAGACCCACAGACACCCACCCAGACCTGGGCCCCGGC	1589
Db	1321	TTGGGACCTGAGGACTCTTTCAGACCCACAGACACCCACCCAGACCTGGGCCCCGGC	1380
Qy	1590	CAGTGTGCGCATCAAAATCCGCAATTCGATGCTGAGGAGGACCAATGAAGATCAGG	1649
Db	1381	CAGTGTGCGCATCAAAATCCGCAATTCGATGCTGAGGAGGACCAATGAAGATCAGG	1440
Qy	1650	GACAGGTTTTGGGACCCCTGGCCAGCGCGACCCCTGACCGACCTCGCTACGCTGTGG	1709
Db	1441	GACAGGTTTTGGGACCCCTGGCCAGCGCGACCCCTGACCGACCTCGCTACGCTGTGG	1500
Qy	1710	GGCGGCTTCGTGTACTGTGAAGACCTGTGTGAGCGGTGACCGCTCCGCTGCTCAGCGGC	1769
Db	1501	GGCGGCTTCGTGTACTGTGAAGACCTGTGTGAGCGGTGACCGCTCCGCTGCTCAGCGGC	1560
Qy	1770	GCCAAACCCCGGGCGGCTCTTACCTGAGCAGATGCCCTATCCGTGCTATGTGGACAC	1829
Db	1561	GCCAAACCCCGGGCGGCTCTTACCTGAGCAGATGCCCTATCCGTGCTATGTGGACAC	1620
Qy	1830	GTGTTCTCTGCTGTGCTGAGCGGCTGCTGCGCTCTTCTGAGCGCTGGCTGATCTTAC	1889
Db	1621	GTGTTCTCTGCTGTGCTGAGCGGCTGCTGCGCTCTTCTGAGCGCTGGCTGATCTTAC	1680
Qy	1890	TCCGTGACACTGACAGTGAAGCGCTGTGTGGGAGAGAGAGACCGCGCTGCGGACACC	1949
Db	1681	TCCGTGACACTGACAGTGAAGCGCTGTGTGGGAGAGAGAGACCGCGCTGCGGACACC	1740
Qy	1950	ATGCGCGGCTGAGGCTCAGCGCGGCTGCTCTGGTGTAGGCTGCTCCTCAGTGGCTC	2009
Db	1741	ATGCGCGGCTGAGGCTCAGCGCGGCTGCTCTGGTGTAGGCTGCTCCTCAGTGGCTC	1800
Qy	2010	GGGCGCTTCTGCTCAGCGCGGCTGCTGCTGCTTCTGCTGCTCAAGCTGGGGACATCTCTC	2069

1801 GGGCCCTCTCTGCTCAGCGCGACACTGCTGTTCTGTTCTCAAGCTGGAGACATCCTC 1860
2070 CCTACAGCCACCGGCGTGGTCTTCTCTCTTTGTCAGCCCTTGGCGTGGCCAGCGTG 2129
1861 CCTACAGCCACCGGCGTGGTCTTCTCTCTTTTGGCAGCCCTTGGCGTGGCCAGCGTG 1920
2130 ACCAGAGCTCTGCTCAGGCGCTTCTTCTCCGCGCCAACTGGCTGGCTGGCGCTGGGC 2189
1921 ACCAGAGCTCTGCTCAGGCGCTTCTTCTCCGCGCCAACTGGCTGGCGCTGGGC 1980
2190 GGCCTGGCCTACTTCTCCCTCTACCTGCCCTAGCTGCTGTGTGTGGCTGGCGGACCGG 2249
1981 GGCCTGGCCTACTTCTCCCTCTACCTGCCCTAGCTGCTGTGTGTGGCTGGCGGACCGG 2040
2250 CTGCCCGCGGTGGCGCGTGGCGGAGCTGCTGTGCGCCCGTGGCTTTGCGCTTGGC 2309
2041 CTGCCCGCGGTGGCGCGTGGCGGAGCTGCTGTGCGCCCGTGGCTTTGCGCTTGGC 2100
2310 TGGAGAGCTGGCTCTGCTGGAGGAGCAGGCGAGGCGCGGAGTGGCAACAGTGGGC 2369
2101 TGGAGAGCTGGCTCTGCTGGAGGAGCAGGCGAGGCGCGGAGTGGCAACAGTGGGC 2160
2370 ACCCGGCTTACGGCAGACGCTTTCAGCCTGGCCCGCAGGTCTCTGGCCCTTCTGCTGCTGGAC 2429
2161 ACCCGGCTTACGGCAGACGCTTTCAGCCTGGCCCGCAGGTCTCTGGCCCTTCTGCTGCTGGAC 2220
2430 GCGGCGCTTACGGCCTCGCCACTGTGTACTGGAAGCTGTGTGCCAGGCCAGTACGGG 2489
2221 GCGGCGCTTACGGCCTCGCCACTGTGTACTGGAAGCTGTGTGCCAGGCCAGTACGGG 2280
2490 ATCCCTGAACCATGGAAATTTCTTTTCGAGGAGCTACTGTTGGGACCTTCGGCCCCC 2549
2281 ATCCCTGAACCATGGAAATTTCTTTTCGAGGAGCTACTGTTGGGACCTTCGGCCCCC 2340
2550 AAGAGTCCAGCCCTTGCCCCACCCCGCTGGACCCCAAGGTGCTGTGTAAGAGGACACCG 2609
2341 AAGAGTCCAGCCCTTGCCCCACCCCGCTGGACCCCAAGGTGCTGTGTAAGAGGACACCG 2400
2610 CCGGCGCTGAGTCTGCGCTATCCGTTTCGAGCTGTGGAGAGCGCTTTCTGGAAGCCG 2669
2401 CCGGCGCTGAGTCTGCGCTCTCCGTTTCGAGCTGTGGAGAGCGCTTTCTGGAAGCCG 2460
2670 CAGCCAGCCCTTGCGGGGCTCAGCCTGGAATTTCTACCGAGGCGCACATCAGCGCTTCTTG 2729
2461 CAGCCAGCCCTTGCGGGGCTCAGCCTGGAATTTCTACCGAGGCGCACATCAGCGCTTCTTG 2520
2730 GGCACAAACGGGCGCGGCAAGAACACCACTGTGTCATCTTGAAGTGGCCTTTCCACCC 2789
2521 GGCACAAACGGGCGCGGCAAGAACCACTGTGTCATCTTGAAGTGGCCTTTCCACCC 2580
2790 AGTGTGGCTCTGCTTTCATCTGGGCGCAGCGTCCGCTCCAGCATGAGCGCGCATCCGG 2849
2581 AGTGTGGCTCTGCTTTCATCTGGGCGCAGCGTCCGCTCCAGCATGAGCGCGCATCCGG 2640
2850 CCCACCTTGCGGCTCTGCTCTCAGTACAACTGCTGTTTGAATGCTGACCTGGAGCGAG 2909
2641 CCCACCTTGCGGCTCTGCTCTCAGTACAACTGCTGTTTGAATGCTGACCTGGAGCGAG 2700
2910 CAGTCTGTGTTTATGGGCGGCTGAAGGGTCTGAGTGGCCTGTAGTGGGCGCCGAGCAG 2969
2701 CAGTCTGTGTTTATGGGCGGCTGAAGGGTCTGAGTGGCCTGTAGTGGGCGCCGAGCAG 2760
2970 GACCGTCTGTCAGGATGTGGGGCTGGTCTCAAGCAGAGTGTGACATTCGCCACCTC 3029
2761 GACCGTCTGTCAGGATGTGGGGCTGGTCTCAAGCAGAGTGTGACATTCGCCACCTC 2820
3030 TCTGTGGGATGCAACCGAAGCTGTCCGTGGCCATTGCTTTGTTGGCGGCTCCCAAGTT 3089
2821 TCTGTGGGATGCAACCGAAGCTGTCCGTGGCCATTGCTTTGTTGGCGGCTCCCAAGTT 2880
3090 GTTATCTCTGACGAGCTACGGCTGGCGTGGATCTCTGTTCCCGCGCGCGTATTGGGAG 3149

2881 GTTATCTCTGACGAGCTACGGCTGGCTGGATCTCTGCTTCCCGCGCGGTATTGGGAG 2940
3150 CTGCTGCTCAAATACCGAGAGGTTCGACGCTGATCTCTTCCACCCACCACTGGATGAG 3209
2941 CTGCTGCTCAAATACCGAGAGGTTCGACGCTGATCTCTTCCACCCACCACTGGATGAG 3000
3210 GCAGAGCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTCTGCTGTGGC 3269
3001 GCAGAGCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTCTGCTGTGGC 3060
3270 TCCCCACTCTTCTGCGCGCTTCACTGGGCTCCGGCTACTTACCTGACGCTGGTGAAGGCC 3329
3061 TCCCCACTCTTCTGCGCGCTTCACTGGGCTCCGGCTACTTACCTGACGCTGGTGAAGGCC 3120
3330 CGCTGCCCTTGAACCAATGAGAGGCTGACACTGACATGAGAGGGCAGTGTGGACAC 3389
3121 CGCTGCCCTTGAACCAATGAGAGGCTGACACTGACATGAGAGGGCAGTGTGGACAC 3180
3390 AGCAGGAAAAGAAATGSCAGCGGCGAGAGTGGCACTCTCTCAGCTTGTGGCC 3449
3181 AGCAGGAAAAGAAATGSCAGCGGCGAGAGTGGCACTCTCTCAGCTTGTGGCC 3240
3450 CTGGTACAGCACTTGGGTGGCGGCGCACGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3509
3241 CTGGTACAGCACTTGGGTGGCGGCGCACGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3300
3510 CTGGTCTGCCCTACAGGGGTGCCATGACGGGAGCTTGGCCACACTCTTCCGAGAGCTA 3569
3301 CTGGTCTGCCCTACAGGGGTGCCATGACGGGAGCTTGGCCACACTCTTCCGAGAGCTA 3360
3570 GACACCGGCTGCGGAGCTGAGGCTCACTGGGCTAGCGGATCTCCGACACAGCTCGAG 3629
3421 GAGATCTTCTGAAGTGGTGGAGGAGTGTGCTGCGGACACAGATATGAGGATGGCAGC 3480
3690 TGGGCGGAGCACTATGCAACGSCATTTGCTGGCTAGACGTAACCTTGGGCTCAAGATG 3749
3481 TGGGCGGAGCACTATGCAACGSCATTTGCTGGCTAGACGTAACCTTGGGCTCAAGATG 3540
3750 CCGCCACAGGAGACGCGCTGGAGAACGGGAAACAGCTGGGTGAGCCCCAGAGACTGAC 3809
3541 CCGCCACAGGAGACGCGCTGGAGAACGGGAAACAGCTGGGTGAGCCCCAGAGACTGAC 3600
3810 CAGGGCTCTGGGCGCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG 3869
3601 CAGGGCTCTGGGCGCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG 3660
3870 CTCAGGCGCTTCTCAAGCGCTTTCTGCTTGGCGCGGAGCGCGCGGCGCTGTTTC 3929
3661 CTCAGGCGCTTCTTCAAGCGCTTTCTGCTTGGCGCGGAGCGCGCGGCGCTGTTTC 3720
3930 GCCCAGATGCTGCTGCTGCTCTTTTGGGCGCTGGCCCTCGTGTTCAGCCTCATCGTG 3989
3721 GCCCAGATGCTGCTGCTGCTCTTTTGGGCGCTGGCCCTCGTGTTCAGCCTCATCGTG 3780
3990 CTTCTTTTGGGCACTACCGCGCTCTGCGGCTCAGTGTCCACCACTGATGATCGGTGTCAGGTG 4049
3781 CTTCTTTTGGGCACTACCGCGCTCTGCGGCTCAGTGTCCACCACTGATGATCGGTGTCAGGTG 3840
4050 TCTTTTCTCAGTGAGAGCGCCCCCAGGGGACCTTGGAGGTGCCCGGCTGCTCGAGGGCGTG 4109
3841 TCTTTTCTCAGTGAGAGCGCCCCCAGGGGACCTTGGAGGTGCCCGGCTGCTCGAGGGCGTG 3900
4110 CTCAGGAGGACGAGCTTGGAGGAGCGCCCAAGTGTGAGCATAGCTCCCAAGGTTCTCGGCA 4169
3901 CTCAGGAGGACGAGCTTGGAGGAGCGCCCAAGTGTGAGCATAGCTCCCAAGGTTCTCGGCA 3960
4170 CCAGAGTTCTCTGCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCT 4229
3961 CCAGAGTTCTCTGCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCT 4020

QY 4230 CCATCCCAAGCCTGCCAGTGTAGCCAGCCCGGTGCCCGCGCTCTCTCCCGACTGCCCG 4289
DB 4021 CCATCCCAAGCCTGCCAGTGTAGCCAGCCCGGTGCCCGCGCTCTCTCCCGACTGCCCG 4080
QY 4290 GCTCAGCTGGTGGTCCCTCCCGCCAGAGGAGTGAACGGCTCTGGGGAGTGGTTAG 4349
DB 4081 GCTCAGCTGGTGGTCCCTCCCGCCAGAGGAGTGAACGGCTCTGGGGAGTGGTTAG 4140
QY 4350 AACCTGACAGCCCGGAACTGTCTGACTTCTGGTCAAGACTCAACCCCGCTGTGTGCG 4409
DB 4141 AACCTGACAGCCCGGAACTGTCTGACTTCTGGTCAAGACTCAACCCCGCTGTGTGCG 4200
QY 4410 CAGGCCCTGAAGACTAAGAGTGGTGAATGAGTCAAGTCAAGAGGTTCTCGCTGGG 4469
DB 4201 CAGGCCCTGAAGACTAAGAGTGGTGAATGAGTCAAGTCAAGAGGTTCTCGCTGGG 4260
QY 4470 GGGCGAGACCCAGGCTCCCTCCCGCCAGAGTGGGGCGCTCAGTGGAGGAGTTGG 4529
DB 4261 GGGCGAGACCCAGGCTCCCTCCCGCCAGAGTGGGGCGCTCAGTGGAGGAGTTGG 4320
QY 4530 GCGCTGCTGAGTCCCTCCCTCCCGCGGCGCTCCAGCCGCTGCTGAAACCTCAAGCC 4589
DB 4321 GCGCTGCTGAGTCCCTCCCTCCCGCGGCGCTCCAGCCGCTGCTGAAACCTCAAGCC 4380
QY 4590 TGGGCTCAAGCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTG 4649
DB 4381 TGGGCTCAAGCTGGACGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTG 4440
QY 4650 CACTCCATGGTGGCTTTGTCAACCGAGCGAGCAACGCAATCTCCGTGCTCACTGCC 4709
DB 4441 CACTCCATGGTGGCTTTGTCAACCGAGCGAGCAACGCAATCTCCGTGCTCACTGCC 4500
QY 4710 CCAGSCCGGCGCGCACCGCACAGCATCACCACTCAACCCCTGAACTCAAC 4769
DB 4501 CCAGSCCGGCGCGCACCGCACAGCATCACCACTCAACCCCTGAACTCAAC 4560
QY 4770 AAGGAGCAGCTGTCTGAGGCTGCACTGATGGCTCTCCGTGGAGCTCTCGTCTCCATC 4829
DB 4561 AAGGAGCAGCTGTCTGAGGCTGCACTGATGGCTCTCCGTGGAGCTCTCGTCTCCATC 4620
QY 4830 TGTGTGCTTTGGCATGCTTTGTCCGGCGAGCTTCACTCTGTGCTCATTTAGGAG 4889
DB 4621 TGTGTGCTTTGGCATGCTTTGTCCGGCGAGCTTCACTCTGTGCTCATTTAGGAG 4680
QY 4890 CGAGTCAACCCGAGCAACGACCTGAGCTCATGGGGGCGCTGTCCCCCACCCTCTACTGG 4949
DB 4681 CGAGTCAACCCGAGCAACGACCTGAGCTCATGGGGGCGCTGTCCCCCACCCTCTACTGG 4740
QY 4950 CTTCGCAACTTTCTCTGGGACATGTGTAACCTTGTGGTCCAGCATGATCGTGTGCTC 5009
DB 4741 CTTCGCAACTTTCTCTGGGACATGTGTAACCTTGTGGTCCAGCATGATCGTGTGCTC 4800
QY 5010 ATCTTTCTGGCTTCCAGAGAGGCAATGTGGCCCTCGCAACCTGCTGCTCTCTG 5069
DB 4801 ATCTTTCTGGCTTCCAGAGAGGCAATGTGGCCCTCGCAACCTGCTGCTCTCTG 4860
QY 5070 CTGTGCTACTGATGCTGGTGGTCAACCGCTCATGATACCCAGCCCTCTCTTC 5129
DB 4861 CTGTGCTACTGATGCTGGTGGTCAACCGCTCATGATACCCAGCCCTCTCTTC 4920
QY 5130 TTCTCCGTGGCCAGCAGACGCTATGTGGTGTCACTGTCATAAACCTCTTTATTTGGCATC 5189
DB 4921 TTCTCCGTGGCCAGCAGACGCTATGTGGTGTCACTGTCATAAACCTCTTTATTTGGCATC 4980
QY 5190 AATGGAAGCATGGCCACCTTTGTGCTGAGCTTTCTCTGATCAAGAGCTCAGAGGTG 5249
DB 4981 AATGGAAGCATGGCCACCTTTGTGCTGAGCTTTCTCTGATCAAGAGCTCAGAGGTG 5040
QY 5250 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGACCTTCTGTTGGGCGGGGGCTC 5309
DB 5041 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGACCTTCTGTTGGGCGGGGGCTT 5100

QY 5310 ATTGACATGTGTCGGAACCAAGCCCATGCTGTATGCTTTGAGCGCTTGGAGACAGGCAG 5369
DB 5101 ATTGACATGTGTCGGAACCAAGCCCATGCTGTATGCTTTGAGCGCTTGGAGACAGGCAG 5160
QY 5370 TTCCAGTCAACCCCTGCGCTGGAGGTGTGCGGAGAACCTCTTGGCCATGCTGATACAG 5429
DB 5161 TTCCAGTCAACCCCTGCGCTGGAGGTGTGCGGAGAACCTCTTGGCCATGCTGATACAG 5220
QY 5430 GGGCCCTCTTCTCTTCTTCACTACTGCTGACGACCGAAGCAAACTCTCTGCCACAG 5489
DB 5221 GGGCCCTCTTCTCTTCTTCACTACTGCTGACGACCGAAGCAAACTCTCTGCCACAG 5280
QY 5490 CCCAGGTGAGGTCTCTGCCACTCTCTGGAGAGAGGAGAGAGTGTAGCCCGTGAACG 5549
DB 5281 CCCAGGTGAGGTCTCTGCCACTCTCTGGAGAGAGGAGAGAGTGTAGCCCGTGAACG 5340
QY 5550 GAGCGGTGTGTCAGAGGACCAACCGAGGGGAGTGTGTTGCTGACAGAACTTGACCAAG 5609
DB 5341 GAGCGGTGTGTCAGAGGACCAACCGAGGGGAGTGTGTTGCTGAGAACTTGACCAAG 5400
QY 5610 GTATACCGTGGGACAGAGATGCCAGCTTGTGACCGCTTGTGCTGGGATTTCCCTCTGT 5669
DB 5401 GTATACCGTGGGACAGAGATGCCAGCTTGTGACCGCTTGTGCTGGGATTTCCCTCTGT 5460
QY 5670 GAGTGTGTTGGCTGCTGGGTGTGAATGGAGCAGGGAAAGCTCCACTTTGCAATGTG 5729
DB 5461 GAGTGTGTTGGCTGCTGGGTGTGAATGGAGCAGGGAAAGCTCCACTTTGCAATGTG 5520
QY 5730 ACGGGGGACATTTGGGCCAGCAGGGCGAGCTGTGCTGGCAGGCCACAGCGTGGCCCG 5789
DB 5521 ACGGGGGACATTTGGGCCAGCAGGGCGAGCTGTGCTGGCAGGCCACAGCGTGGCCCG 5580
QY 5790 GAAACCAAGTGTGCGCACTCAGCATGGATGACTGCGCTCAATCCGATGCCATCTTTGAG 5849
DB 5581 GAAACCAAGTGTGCGCACTCAGCATGGATGACTGCGCTCAATCCGATGCCATCTTTGAG 5640
QY 5850 CTGTGACGGGCCCGAGCACTCTGAGCTGTGTCGCGCTCTGCGGCTGTCCCGAGGCC 5909
DB 5641 CTGTGACGGGCCCGAGCACTCTGAGCTGTGTCGCGCTCTGCGGCTGTCCCGAGGCC 5700
QY 5910 CAGTGTGCCAGACGGCTGGCTCAGGCTGCGGCTGTGGGACTCTCATGTATCCGAGC 5969
DB 5701 CAGTGTGCCAGACGGCTGGCTCAGGCTGCGGCTGTGGGACTCTCATGTATCCGAGC 5760
QY 5970 CGGCTCGCAGGCACTCAGCGAGGGGAAACAAACGCAAGCTGGCGAGCCCTGCGCTG 6029
DB 5761 CGGCTCGCAGGCACTCAGCGAGGGGAAACAAACGCAAGCTGGCGAGCCCTGCGCTG 5820
QY 6030 GTTGGGACCCAGCCGCTGTGTTTCTGGAACGAGCCGACACAGGAGTGGACCCAGCGCG 6089
DB 5821 GTTGGGACCCAGCCGCTGTGTTTCTGGAACGAGCCGACACAGGAGTGGACCCAGCGCG 5880
QY 6090 CGGCGCTTCTTTGGAAACAGCTTTTGGCCGTGTGGTGGGAGGGCGCTTCAGTATGCTC 6149
DB 5881 CGGCGCTTCTTTGGAAACAGCTTTTGGCCGTGTGGTGGGAGGGCGCTTCAGTATGCTC 5940
QY 6150 ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGTAAT 6209
DB 5941 ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGTAAT 6000
QY 6210 GGGCGGTTCCGCTGCTGGGAGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCAAC 6269
DB 6001 GGGCGGTTCCGCTGCTGGGAGCCGCAACATCTCAAGGGCAGATTCGCGCGGGTCAAC 6060
QY 6270 ACACTGACCTTGGGTGCGCGCGCAAGTTCACAGCCGAGCGGCTTGTGTGGCGGCC 6329
DB 6061 ACACTGACCTTGGGTGCGCGCGCAAGTTCACAGCGCGAGCGGCTTGTGTGGCGGCC 6120
QY 6330 GAGTTCCTTGGTGGAGCTGCGGAGGCAATGAGGCGCCCTGCGCTTCCAGCTGCG 6389
DB 6121 GAGTTCCTTGGTGGAGCTGCGGAGGCAATGAGGCGCCCTGCGCTTCCAGCTGCG 6180
QY 6390 CCGGAGGGCGCTGCGCCCTGCGCGCTCTTTTGGAGAGCTGGGGGTGCAAGCGCGCAGAG 6449

[illegible]

RESULT 9

```

US-10-182-006-1
; Sequence 1, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23
; PRIORITY APPLICATION NUMBER: PCT/US01/02191
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6509)
; OTHER INFORMATION: Any
US-10-182-006-1

```

	Query Match	95.2%	Score 6478.4;	DB 16;	Length 6522;
	Best Local Similarity	99.7%;	Pred. No. 0;		
	Matches 6501; Conservative	0;	Mismatches	17; Indels	3; Gaps 1;
Qy	210	ATGCGCTTCTGGACACAGCTGATGCTGCTCTGTGGAGAATTTCATGTAATGCCCGAGA	269		
Dd	1	ATGCGCTTCTGGACACAGCTGATGCTGCTCTGTGGAGAATTTCATGTAATGCCCGAGA	60		
Qy	270	CAGCCGGTCCAGCTCCTGGTCGAATTGCTGTGGCCTCTCTTCCTTTCTTTCATCTGGTG	329		
Dd	61	CAGCCGGTCCAGCTCCTGGTCGAATTGCTGTGGCCTCTCTTCCTTTCTTTCATCTGGTG	120		
Qy	330	GCTGTTTCGCACATCCCAACCGCCCTCGAGACAACATGAATGCCACTTCCCAAACAAGCCA	389		
Dd	121	GCTGTTTCGCACATCCCAACCGCCCTCGAGACAACATGAATGCCACTTCCCAAACAAGCCA	180		

Qy	390	CTGCCATCGCGGGCACCGTGGCCCTGGCTCCAGGGTCTCATCTGTGAATGTGAACAACACC	449
Db	181	CTGCCATCGCGGGCACCGTGGCCCTGGCTCCAGGGTCTCATCTGTGAATGTGAACAACACC	240
Qy	450	TGCTTTCCGCAGCTGACACCGGCGGAGGAGCCCGGGCGCTGTAGACAACTTCAACGACTCC	509
Db	241	TGCTTTCCGCAGCTGACACCGGCGGAGGAGCCCGGGCGCTGTAGACAACTTCAACGACTCC	300
Qy	510	CTGCTTCCCGGCTGTAGCCGATGCGCAGTGTGTGTGGAGGGGCCAGTGCGCCACAGG	569
Db	301	CTGGTCTCCCGGCTGTAGCCGATGCGCAGTGTGTGTGGAGGGGCCAGTGCGCCACAGG	360
Qy	570	ACGCTGCTGGCTTAGGGAAGCTGATGCGCACGCTGTAGGGCTGCAACGACGACGCGGCCAG	629
Db	361	ACGCTGCTGGCTTAGGGAAGCTGATGCGCACGCTGTAGGGCTGCAACGACGACGCGGCCAG	420
Qy	630	CCTCAACCAACCAAGCAGTCTCCACTGTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	689
Db	421	CCTCAACCAACCAAGCAGTCTCCACTGTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	480
Qy	690	ACGTCACTGTGCGCACGGAAATCCCTGGGGTGTGGCACTGGGCCCAAGCCACGAGAGCCCTTG	749
Db	481	ACGTCACTGTGCGCACGGAAATCCCTGGGGTGTGGCACTGGGCCCAAGCCACGAGAGCCCTTG	540
Qy	750	CACAGCTTGTGGAGCCGCTGAGGACCTGGCCCCAGGAGCTCTCTGGCGTGTGCCAGGCTG	809
Db	541	CACAGCTTGTGGAGCCGCTGAGGACCTGGCCCCAGGAGCTCTCTGGCGTGTGCCAGGCTG	600
Qy	810	GTGAGCTTCGGGCACATGCTGCAGAGACCCCGAGGAGACAGCGGCCGCCCTCTGGAGTTGCTG	869
Db	601	GTGAGCTTCGGGCACATGCTGCAGAGACCCCGAGGAGACAGCGGCCGCCCTCTGGAGTTGCTG	660
Qy	870	TCAGAGCCCTCTGCAAGTGTACGGGACCTTAGCAGCACAGTGTGGCGCCCTCCCTCAACTGG	929
Db	661	TCAGAGCCCTCTGCAAGTGTACGGGAGCTTAGCAGCACAGTGTGGCGCCCTCCCTCAACTGG	720
Qy	930	TACGAGCTAGTGACCTGTATGGAGCTGTGTGGGCGAGGAGCAGAAATCCGCCCTGCCAGAC	989
Db	721	TACGAGCTAGTGACCTGTATGGAGCTGTGTGGGCGAGGAGCAGAAATCCGCCCTGCCAGAC	780
Qy	990	AGACGCTGAGCCCCCGCTGCTCGGAGCTGATTGGAGCCCTTGACAGCCACCCGCTGTCC	1049
Db	781	AGCAGCTGAGCCCCCGCTGCTCGGAGCTGATTGGAGCCCTTGACAGCCACCCGCTGTCC	840
Qy	1050	CGCTGCTCTGGAGACGCTCTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGCACACAGAT	1109
Db	841	CGCTGCTCTGGAGAGCGCTCTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGCACACAGAT	900
Qy	1110	ACACCTTTTACCGGAAGCTCATGGCCAGGTTGAACCGGACCTTTCGAGGAGCTCACCCCTG	1169
Db	901	ACACCTTTTACCGGAAGCTCATGGCCAGGTTGAACCGGACCTTTCGAGGAGCTCACCCCTG	960
Qy	1170	CTGAGGATGTCCGGAGAGTGTGGGAGATGCTGGGACCCCGGATCTTTCACCTTATGAAC	1229
Db	961	CTGAGGATGTCCGGAGAGTGTGGGAGATGCTGGGACCCCGGATCTTTCACCTTATGAAC	1020
Qy	1230	GACAGTTCCTAACTGTGGCCTGCTGACGCGCTCCTTCAGATGACAGATGAAGGAAGAAGG	1289
Db	1021	GACAGTTCCTAACTGTGGCCTGCTGACGCGCTCCTTCAGATGACAGATGAAGGAAGAAGG	1080
Qy	1290	CAGCCACAGACTGGAGGCCGGGACCATGGAGGCCCTGCGATCTTTCGAGCCCTGGG	1349
Db	1081	CAGCCACAGACTGGAGGCCGGGACCATGGAGGCCCTGCGATCTTTCGAGCCCTGGG	1140
Qy	1350	AGCGGTGGCTTACAGCTGGGACGACGACAGCTGTATGTGGGGCACTGTGTGGGCAAGCTG	1409
Db	1141	AGCGGTGGCTTACAGCTGGGACGACGACAGCTGTATGTGGGGCACTGTGTGGGCAAGCTG	1200
Qy	1410	GSCGAGTGCAGAGTGCCTGCTCTTGGACAAGCTGTGAGCGGCACCCCTCAGAGGCGACC	1469
Db	1201	GSCGAGTGCAGAGTGCCTGCTCTTGGACAAGCTGTGAGCGGCACCCCTCAGAGGCGACC	1260
Qy	1470	CTGGTGTGCGGGGCCCTTGCAACTGCTCGCGGAACATCGATTTCTGGGCGCGGCTGCTTTC	1529

QY 3690 TGGGGCAGCACCCTATGSCAAGGCAATTGCTGGCTAGACGTAAACCTTGGGCTCAAGATG 3749
Db 3481 TGGGGCAGCACCCTATGSCAAGGCAATTGCTGGCTAGACGTAAACCTTGGGCTCAAGATG 3540
QY 3750 CCGCCACAGAGACAGCGCTGGAGAAACGGGGAACCAAGCTGGGTAGAGAGAGCTGAC 3809
Db 3541 CCGCCACAGAGACAGCGCTGGAGAAACGGGGAACCAAGCTGGGTAGAGAGAGCTGAC 3600
QY 3810 CAGGGCTCTGGGCGAGACGCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCCACGAG 3869
Db 3601 CAGGGCTCTGGGCGAGACGCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCCACGAG 3660
QY 3870 CTCAGGCGCTGCTCTCAAGCGCTTCTGCTTGGCGCGCAGCGCGCGCGCTGCTTC 3929
Db 3661 CTCAGGCGCTGCTCTCAAGCGCTTCTGCTTGGCGCGCAGCGCGCGCGCTGCTTC 3720
QY 3930 GCCCAGATCGTGTGCTGCGCTCTTGTGGGCTGGCGCTCGTGTTCAGCCTCATCGTG 3989
Db 3721 GCCCAGATCGTGTGCTGCGCTCTTGTGGGCTGGCGCTCGTGTTCAGCCTCATCGTG 3780
QY 3990 CTTCTTTTGGGCACTACCGGCTCTGGGCTCAGTCCCAACCATGTACGGTGTCAAGGT 4049
Db 3781 CTTCTTTTGGGCACTACCGGCTCTGGGCTCAGTCCCAACCATGTACGGTGTCAAGGT 3840
QY 4050 TCCCTTCTCAGTGAGGACGCCAGGGGACCTGGAAGCTGCGCGCTGCTCGAGGCGCTG 4109
Db 3841 TCCCTTCTCAGTGAGGACGCCAGGGGACCTGGAAGCTGCGCGCTGCTCGAGGCGCTG 3900
QY 4110 CTGAGGAGCAGGACTGGAGGAGCCCCCAGTCAGCATAGCTCCACAGGTTCTCGGCA 4169
Db 3901 CTGAGGAGCAGGACTGGAGGAGCCCCCAGTCAGCATAGCTCCACAGGTTCTCGGCA 3960
QY 4170 CAGAAAGTTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAATGGACCCAGAGTCT 4229
Db 3961 CAGAAAGTTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAATGGACCCAGAGTCT 4020
QY 4230 CCATCCCGAGCTGCCAGTGTAGCCAGCCGGTGGCGGCGCTGCTGCCAGCTGCCG 4289
Db 4021 CCATCCCGAGCTGCCAGTGTAGCCAGCCGGTGGCGGCGCTGCTGCCAGCTGCCG 4080
QY 4290 GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGAGTACCGGCTCTGGGGAAGTGGTTCA 4349
Db 4081 GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGAGTACCGGCTCTGGGGAAGTGGTTCA 4140
QY 4350 AACCTGACAGGCGGGAACCTGTCTGACTTCTGGTCAAGACCTTACCGCGCTTGGTGGC 4409
Db 4141 AACCTGACAGGCGGGAACCTGTCTGACTTCTGGTCAAGACCTTACCGCGCTTGGTGGC 4200
QY 4410 CAGGGCTGAAGACTGAAGTGGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGG 4469
Db 4201 CAGGGCTGAAGACTGAAGTGGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGG 4260
QY 4470 GGGCGAGACCCAGGCTGCGCTCGGGCCAAAGTGGGGCGCTCAGTGGAGGAGTGGTGG 4529
Db 4261 GGGCGAGACCCAGGCTGCGCTCGGGCCAAAGTGGGGCGCTCAGTGGAGGAGTGGTGG 4320
QY 4530 GCCTGCTGAGTCCCCCTGCTGGCGGGGCCCTCGA CCGTGTCTGAAAAAACCCTCAAGCC 4589
Db 4321 GCCTGCTGAGTCCCCCTGCTGGCGGGGCCCTCGA CCGTGTCTGAAAAAACCCTCAAGCC 4380
QY 4590 TGGGCTCAGAGCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGG 4649
Db 4381 TGGGCTCAGAGCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGG 4440
QY 4650 CACTCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCC 4709
Db 4441 CACTCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCC 4500
QY 4710 CAGGCGCGCGCGCCACGCCCCAAGCATCAACACTCAACCAACCCCTTTGAACCTCAAC 4769
Db 4501 CAGGCGCGCGCGCCACGCCCCAAGCATCAACACTCAACCAACCCCTTTGAACCTCAAC 4560

QY 4770 AAGGAGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGACGCTCTCTGCTCCATC 4829
Db 4561 AAGGAGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGACGCTCTCTGCTCCATC 4620
QY 4830 TGTGTGGTCTTTTGGCCATGTCTTTGTCCGGGCGAGTTTCACTTGTCTCATTTGAGGAG 4889
Db 4621 TGTGTGGTCTTTTGGCCATGTCTTTGTCCGGGCGAGTTTCACTTGTCTCATTTGAGGAG 4680
QY 4890 CGAGTCAACCGAGCCAAAGCACCTGAGCTCATGGGGGGCTGTCCCGCACCTCTACTGG 4949
Db 4681 CGAGTCAACCGAGCCAAAGCACCTGAGCTCATGGGGGGCTGTCCCGCACCTCTACTGG 4740
QY 4950 CTTGGCAAATTTCTCTGGGACATGTGTAATCTACTTGTGTGCCAGCATGTCGTGGTGTCTC 5009
Db 4741 CTTGGCAAATTTCTCTGGGACATGTGTAATCTACTTGTGTGCCAGCATGTCGTGGTGTCTC 4800
QY 5010 ATCTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCTTGCACCACTGTCTCTCTCTTC 5069
Db 4801 ATCTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCTTGCACCACTGTCTCTCTCTTC 4860
QY 5070 CTGTTGCTACTACTGTATGGCTGATCACAACCGCTCATGTACCCAGCCTCTCTCTTC 5129
Db 4861 CTGTTGCTACTACTGTATGGCTGATCACAACCGCTCATGTACCCAGCCTCTCTCTTC 4920
QY 5130 TTCTCGTGCCTCAGCAGCAGCTATGTGTGTGCTCACCTGCATAAACCTCTTTATTGGCATC 5189
Db 4921 TTCTCGTGCCTCAGCAGCAGCTATGTGTGTGCTCACCTGCATAAACCTCTTTATTGGCATC 4980
QY 5190 AATGGAAGCATGGCCACCTTTTGTGTGAGTCTTCTCTGATCAGAAAGCTGCAAGGAGTG 5249
Db 4981 AATGGAAGCATGGCCACCTTTTGTGTGAGTCTTCTCTGATCAGAAAGCTGCAAGGAGTG 5040
QY 5250 AGCGGATCTTGAACAGGCTCTCTTATTTCCCGCATCTTCTGCTGGCGCGGGGGCTC 5309
Db 5041 AGCGGATCTTGAACAGGCTCTCTTATTTCCCGCATCTTCTGCTGGCGCGGGGGCTC 5100
QY 5310 ATTGACATGCTGGGAAACAGGCACTGGCTGTATGCTTTGAGCGCTTGGGAGACAGGAG 5369
Db 5101 ATTGACATGCTGGGAAACAGGCACTGGCTGTATGCTTTGAGCGCTTGGGAGACAGGAG 5160
QY 5370 TTCCAGTCAACCTTGGCTGGGAGTGGTGGCGAAGAACTCTTTGGCCATTTGATGATACAG 5429
Db 5161 TTCCAGTCAACCTTGGCTGGGAGTGGTGGCGAAGAACTCTTTGGCCATTTGATGATACAG 5220
QY 5430 GGGCCCTCTTCTCTTCTCACACTCTGCTGACAGACCGAAGCCAACTCTCTGCCACAG 5489
Db 5221 GGGCCCTCTTCTCTTCTCACACTCTGCTGACAGACCGAAGCCAACTCTCTGCCACAG 5280
QY 5490 CCGAGGCTGAGGCTCTCTGCACTCTCTGGGAGAGGAGAGGATGAGCCCTGAAACGG 5549
Db 5281 CCGAGGCTGAGGCTCTCTGCACTCTCTGGGAGAGGAGAGGATGAGCCCTGAAACGG 5340
QY 5550 GAGCGGCTGGTCCAAGGAGCCACCCAGGGGATGTGTGTGTGAGGAACTTTGACCAAG 5609
Db 5341 GAGCGGCTGGTCCAAGGAGCCACCCAGGGGATGTGTGTGTGAGGAACTTTGACCAAG 5400
QY 5610 GTATACCTGGGAGAGGATGCGAGCTGTGACCGCTTGTGCTGGGATTTCCCGCTGT 5669
Db 5401 GTATACCTGGGAGAGGATGCGAGCTGTGACCGCTTGTGCTGGGATTTCCCGCTGT 5460
QY 5670 GAGTGTTTGGGCTGTCTGGGTGGAATGGAGAGGAGAGGATGAGCCCTGATGGTG 5729
Db 5461 GAGTGTTTGGGCTGTCTGGGTGGAATGGAGAGGAGAGGATGAGCCCTGATGGTG 5520
QY 5730 ACGGGGAGACATTTGGCCAGAGGCGGAGGCTGTGTGTGTGAGGCAACAGCGTGGCCGG 5789
Db 5521 ACGGGGAGACATTTGGCCAGAGGCGGAGGCTGTGTGTGTGAGGCAACAGCGTGGCCGG 5580
QY 5790 GAACCCAGTGTGGCGCACCTCAGCATGGGATCTGCTCCCTCAATCCCATCTTTTGG 5849
Db 5581 GAACCCAGTGTGGCGCACCTCAGCATGGGATCTGCTCCCTCAATCCCATCTTTTGG 5640
QY 5850 CTGCTGACGGGCGCGAGCACTTGGAGCTGTCTTGGCGGCTGTGCGGAGGCTGTCCCGAGGCC 5909

835	CGCCTGCTCTGGAGACGCGCTGAAGCCCTCTGATCTCTGGGAAGCTACTCTTTTGGACCAAGAT	894
Qy	ACACCTTTTACCCCGAAGCTCATGGCCAGGTGAACCGGACCTTTTCAGGAGAGCTCACCCCTG	1169
Db	ACACCTTTTACCCCGAAGCTCATGGCCAGGTGAACCGGACCTTTTCAGGAGAGCTCACCCCTG	954
Qy	CTGAGGAGTGTCCGGAGGTTGGGAGATGCTGGGACCCCGGATCTTCACTTTCATGAAC	1329
Db	CTGAGGAGTGTCCGGAGGTTGGGAGATGCTGGGACCCCGGATCTTCACTTTCATGAAC	1014
Qy	GACAGTTCCAATGTGGCCATGCTGCAGGGCTCTCTGCAGATGCAGGATGAAGNAGNAGG	1289
Db	GACAGTTCCAATGTGGCCATGCTGCAGGGCTCTCTGCAGATGCAGGATGAAGNAGNAGG	1074
Qy	CAGCCACAGACTGGAGGCGGGACCACTAGGAGCCCTGCGATCTCTTCTGGACCCCTGGG	1349
Db	CAGCCACAGACTGGAGGCGGGACCACTAGGAGCCCTGCGATCTCTTCTGGACCCCTGGG	1134
Qy	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGGCACCTGTGGGACCGCTG	1409
Db	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGGCACCTGTGTGGGACCGCTG	1194
Qy	GGCCGAGTGACGGAGTGCCTTCTTGGACAAAGCTGAGGGGGCACCCCTCAGAGGACGCC	1469
Db	GGCCGAGTGACGGAGTGCCTTCTTGGACAAAGCTGAGGGGGCACCCCTCAGAGGACGCC	1254
Qy	CTGTGTCTGGGGCCCTCGAACCTGCTCGCGAACATCGATTCTTGGGCCGCGCTGCTCTTC	1529
Db	CTGTGTCTGGGGCCCTCGAACCTGCTCGCGAACATCGATTCTTGGGCCGCGCTGCTCTTC	1314
Qy	TTGGGACCTGAGGACTCTTTCAGACCCACAGACGACCCAAACCCAGACCTTGGGCCCGCGC	1589
Db	TTGGGACCTGAGGACTCTTTCAGACCCACAGACGACCCAAACCCAGACCTTGGGCCCGCGC	1374
Qy	CAGGTGGCATCAAATTCGGATGGAATTAAGCTGTGTCAGAGGACCAATAGATCAGG	1649
Db	CAGGTGGCATCAAATTCGGATGGAATTAAGCTGTGTCAGAGGACCAATAGATCAGG	1434
Qy	GACAGGTTTTGGGACCTTGCCAGCCGCGGACCCCTGACCGACCTGCGCTACGTGTGG	1709
Db	GACAGGTTTTGGGACCTTGCCAGCCGCGGACCCCTGACCGACCTGCGCTACGTGTGG	1494
Qy	GGCGGCTTCTGTACTGCAAGACTTGGTGGAGCGTGCAGCCGCTTCGCGTGTCTCAGCGGC	1769
Db	GGCGGCTTCTGTACTGCAAGACTTGGTGGAGCGTGCAGCCGCTTCGCGTGTCTCAGCGGC	1554
Qy	GCCAAACCCCGGGCGGGCTCTACTGGAGCAGATGCCCTATCCGTGCTATGTGGACGAC	1829
Db	GCCAAACCCCGGGCGGGCTCTACTGGAGCAGATGCCCTATCCGTGCTATGTGGACGAC	1614
Qy	GTGTTCCTGGTGTGCTGAGCCGCTGCTGCGCTCTTCTCTGACGCTGGCTCGATCTAC	1889
Db	GTGTTCCTGGTGTGCTGAGCCGCTGCTGCGCTCTTCTCTGACGCTGGCTCGATCTAC	1674
Qy	TCCGTGACCTGACAGTGAAGGCCGCTGGTGGGAGAGGAGACCGCGCTTCGCGGACAC	1949
Db	TCCGTGACCTGACAGTGAAGGCCGCTGGTGGGAGAGGAGACCGCGCTTCGCGGACAC	1734
Qy	ATGCGGCGCAATGGGGCTCAGCCGCGGGTGTCTGGCTAGGCTGGTTCTCAGCTGCGCTC	2009
Db	ATGCGGCGCAATGGGGCTCAGCCGCGGGTGTCTGGCTAGGCTGGTTCTCAGCTGCGCTC	1794
Qy	GGGCGCTTCTGCTCAGGCGCGCTGCTGGTTCTGGTGTCTCAGCTGGGGGACATCCTC	2069
Db	GGGCGCTTCTGCTCAGGCGCGCTGCTGGTTCTGGTGTCTCAGCTGGGGGACATCCTC	1854
Qy	CCCTACAGCCACCCGGGCGTGTCTTCTGTCTTGGCAGCCTTCGCGGTGGCCACGGTG	2129
Db	CCCTACAGCCACCCGGGCGTGTCTTCTGTCTTGGCAGCCTTCGCGGTGGCCACGGTG	1914
Qy	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCGCGCCAACTTGGCTGCGGCTGCGGC	2189
Db	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCGCGCCAACTTGGCTGCGGCTGCGGC	1974

Qy	2190	GGCTGGCCTTACTTCCCTCTACTTGCCTACGTGCTGTGCTGTGGCTGTGGCTGGCGGACCGG	2249
Db	1975	GGCTGGCCTACTTCTCCCTCTACCTGCCCTACGTGCTGTGCTGTGGCTTGGCGGGACCGG	2034
Qy	2250	CTGCCCGGGTGGCGCGCTGGCGGAGCGCTGCTGTCGCCCTGTGGCCCTTGGCTTGGCTTCGGC	2309
Db	2035	CTGCCCGGGTGGCGCGCTGGCGGAGCGCTGCTGTCGCCCTGTGGCCCTTGGCTTGGCTTCGGC	2094
Qy	2310	TGCGAGAGCTGCTCTGCTGGAGGAGCAGGGCGAGGGCGCGAGTGTGGCAACAACGTGGCG	2369
Db	2095	TGCGAGAGCTGCTCTGCTGGAGGAGCAGGGCGAGGGCGCGAGTGTGGCAACAACGTGGCG	2154
Qy	2370	ACCGGGCTACGGCAGACGTCTTTCAGCTTGGCCCCAGGTCTCTTGGCCCTTCTGCTGCTGGAC	2429
Db	2155	ACCGGGCTACGGCAGACGTCTTTCAGCTTGGCCCCAGGTCTCTTGGCCCTTCTGCTGCTGGAC	2214
Qy	2430	GGCGCGCTACGGCGCTTGGCCACTGGTACTTGGAGAGCTGTGTGCCAGGCCAGTACGGG	2489
Db	2215	GGCGCGCTACGGCGCTTGGCCACTGGTACTTGGAGAGCTGTGTGCCAGGCCAGTACGGG	2274
Qy	2490	ATCCCTGAACCATGGAATTTTCTTTTCGGAGAGCTACTGTGTGGAGCTCTCGGCCCGCC	2549
Db	2275	ATCCCTGAACCATGGAATTTTCTTTTCGGAGAGCTACTGTGTGGAGCTCTCGGCCCGCC	2334
Qy	2550	AAGAGTCCAGCCCTTGGCCCCACCCCGCTGGACCCCAAAGTGTGTGTAGAGGCAACCG	2609
Db	2335	AAGAGTCCAGCCCTTGGCCCCACCCCGCTGGACCCCAAAGTGTGTGTAGAGGCAACCG	2394
Qy	2610	CCGGCGCTGAGTCTGTGGCGTATCGTTTCGAGCTGTGAGAGAGCGCTTCTTGGAGGCCG	2669
Db	2395	CCGGCGCTGAGTCTGTGGCGTCTCGTTTCGAGCTGTGAGAGAGCGCTTCTTGGAGGCCG	2454
Qy	2670	CAGCAGCCCTGGGGGCTCAGCTGGACTTCTACCCAGGSCACATCACCGCTTCTCTG	2729
Db	2455	CAGCAGCCCTGGGGGCTCAGCTGGACTTCTACCCAGGSCACATCACCGCTTCTCTG	2514
Qy	2730	GGCCACAAAGGGGCGCGCAAGACACACCGCTGTCCATCTTGTAGTGGCTCTTCCCAACC	2789
Db	2515	GGCCACAGCGGGCGCGCAAGACACACCGCTGTCCATCTTGTAGTGGCTCTTCCCAACC	2574
Qy	2790	AGTGTGGCTCTGCTTTCATCTTGGGCGACGAGCTGCGCTTCAGAGTGGCGCGCATCCGG	2849
Db	2575	AGTGTGGCTCTGCTTTCATCTTGGGCGACGAGCTGCGCTTCAGAGTGGCGCGCATCCGG	2634
Qy	2850	CCCACTGGGCGTGTCTCTCAGTACAACGTGCTGTTTGCATGCTGACCTGGAGCGAG	2909
Db	2635	CCCACTGGGCGTGTCTCTCAGTACAACGTGCTGTTTGCATGCTGACCTGGAGCGAG	2694
Qy	2910	CAGCTCTGTTCTATGGCGGCTGAAGGTTCTGAGTGGCCGCTGTGTTGGGCCCGCGAGCAG	2969
Db	2695	CAGCTCTGTTCTATGGCGGCTGAAGGTTCTGAGTGGCCGCTGTGTTGGGCCCGCGAGCAG	2754
Qy	2970	GACCGTCTGTCAGAGATGTGGGCTGGTCTCAAAGCAGAGTGTGCAGATCTGCGCACCTC	3029
Db	2755	GACCGTCTAATGAGAGATGTGGGCTGGTCTCAAAGCAGAGTGTGCAGATCTGCGCACCTC	2814
Qy	3030	TCTGGTGGGATGCAACGGAGCTGTCGTGGCCATTTGCCCTTGTGGGCGGCTCCCAAGTT	3089
Db	2815	TCTGGTGGGATGCAACGGAGAGCTGTCTGTTGGCCATTTGCCCTTGTGGGCGGCTCCCAAGTT	2874
Qy	3090	GTTATCTGACAGAGCTACGGCTGGCGTGGATCTCTGTTTCCCGCCCGCGGTATTTGGGAG	3149
Db	2875	GTTATCTGACAGAGCTACGGCTGGCGTGGATCTCTGTTTCCCGCCCGCGGTATTTGGGAG	2934
Qy	3150	CTGCTGCTCAAAATACGAGAGGTTCGACGCTGATCTCTCCACCAACCTGATGAG	3209
Db	2935	CTGCTGCTCAAAATACGAGAGGTTCGACGCTGATCTCTCCACCAACCTGATGAG	2994
Qy	3210	GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGCTGTGGC	3269
Db	2995	GCAGAGCTGTGGGAGACCGTGTGGCGGTGGCAGGTGGCGCTGTGCTGCTGTGGC	3054

QY	3270	TCCCACTTTCCTGCGCCGTCACTTG6GCTCGGGCTACTACCTGA	CCTG6GCTCGGGCTACTACCTGA	CGCTGGTGAAGGCC	3329
DB	3055	TCCCCA	CTTTCTGCGCGCTCACCTG6GCTCGGGCTACTACCTGA	CGCTGGTGAAGGCC	3114
QY	3330	CQCCTGCCCCTGACCACAAATGAGAAGCTGACA	CTGACTGACATGAGGAGCAGTGTGGACAC		3389
DB	3115	CGCTGCCCCTGACCA	CNAATGAGAAGCTGACA	CTGACTGAGGAGGAGCATGTGGACAC	3174
QY	3390	AGGCAGGAAAAAGAAATGGCAGCCAGGGCAGCAGAGT	TCGGCACTCCTCAGCTGCTGGCC		3449
DB	3175	AGGCAGGAAAAAGAAATGGCAGCCAGGGCAGCAGAGT	TCGGCACTCCTCAGCTGCTGGCC		3234
QY	3450	CTGGTA	CAGCACTGGGTGCCCGGGGCACGGCTGTGTGAGAGAGTGC	CCACACGAGCTGGT	3509
DB	3235	CTGGTACAGCACTGGGTGCCCGGGGCACGGCTGTGTGAGAGAGT	TCGCCACACGAGCTGGT		3294
QY	3510	CTGGTGTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACAC	CTTCCTCCGAGACTA		3569
DB	3295	CTGGTGTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACAC	CTTCCTCCGAGACTA		3354
QY	3570	GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACAC	CCAGCCTCGAG		3629
DB	3355	GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACAC	CCAGCCTCGAG		3414
QY	3630	GAGATCTTCGTAAGGTGGTGGAGGAGTGTCTCGGACACAGATATG	AGAGATGACGAC		3689
DB	3415	GAGATCTTCGTAAGGTGGTGGAGGAGTGTCTCGGACACAGATATG	AGAGATGACGAC		3474
QY	3690	TGCGGGCAGCACTATGACAGGCATTGCTGGCCTAGACGTAC	CCCTGCGGCTCAAGATG		3749
DB	3475	TGCGGGCAGCACTATGACAGGCATTGCTGGCCTAGACGTAC	CCCTGCGGCTCAAGATG		3534
QY	3750	CCGCCACAGGACAGCGGCTGGAGAACGGGGAAACGAGCTGGGT	CAGCCCGCAGAGACTGAC		3809
DB	3535	CCGCCACAGGACAGCGGCTGGAGAACGGGGAAACGAGCTGGGT	CAGCCCGCAGAGACTGAC		3594
QY	3810	CAGGGCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTGGG	CACTGACCCCGCAGCAG		3869
DB	3595	CAGGGCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTGGG	CACTGACCCCGCAGCAG		3654
QY	3870	CTCCAGGCGGCTTCTCAGAGCGCTTCTGCTGCGCGCGCAGCGCG	CGGCGGCTGTC		3929
DB	3655	CTCCAGGCGGCTTCTCAGAGCGCTTCTGCTGCGCGCGCAGCGCG	CGGCGGCTGTC		3714
QY	3930	GCCCAGATCTGCTGCCCTCTTTGTGGGCTGCGCCCTCGTGT	TACGCTCATCTGCTG		3989
DB	3715	GCCCAGATCTGCTGCCCTCTTTGTGGGCTGCGCCCTCGTGT	TACGCTCATCTGCTG		3774
QY	3990	CCTCCTTTGGGCACTACCGGGCTCTGCGGCTCAGTCCCA	CCATATGATCGGTCTCAGGTG		4049
DB	3775	CCTCCTTTGGGCACTACCGGGCTCTGCGGCTCAGTCCCA	CCATATGATCGGTCTCAGGTG		3834
QY	4050	TCCTTCTCAGTGAGGAGCGCCACAGGGAGCCCTGGACGTG	CCCGGCTGCTCGAGGCGCTG		4109
DB	3835	TCCTTCTCAGTGAGGAGCGCCACAGGGAGCCCTGGACGTG	CCCGGCTGCTCGAGGCGCTG		3894
QY	4110	CTCAGGAGGACAGACTGGAGGAGCCCGCAGTGAGCATAGCT	CCACAGGTTCTCGGCA		4169
DB	3895	CTCAGGAGGACAGACTGGAGGAGCCCGCAGTGAGCATAGCT	CCACAGGTTCTCGGCA		3954
QY	4170	CCAGAGTTCTGCTGAGTGGCCAAAGTCTTGGCCAGTGGCA	CTTGGACCCCGAGTCT		4229
DB	3955	CCAGAGTTCTGCTGAGTGGCCAAAGTCTTGGCCAGTGGCA	CTTGGACCCCGAGTCT		4014
QY	4230	CCATCCCGAGGCTGCGAGTGA	CCGCGGCTGCTGCTGCCGCTGCTGCTGCCG		4289
DB	4015	CCATCCCGAGGCTGCGAGTGA	CCGCGGCTGCTGCTGCCGCTGCTGCTGCCG		4074
QY	4290	GCTCGAGCTGTGTCCCTCGCGCCACGAGTAGACGGGCTCT	CGGGAGTGGTTTCA		4349
DB	4075	GCTCGAGCTGTGTCCCTCGCGCCACGAGTAGACGGGCTCT	CGGGAGTGGTTTCA		4134
QY	4350	AACCTGACGGCGGGAACCTGTCTGACTTCTCGGTCAAGA	CACTACCGCGGCTGTGTGCGC		4409

Db	4135	AACCTGACAGCGCGGAACCTGTCTGACTTCTCTGGTCAAGACCTACCGCGCCTGGTGC	4194
Qy	4410	CAGGGCCTGAAGACTAAGAGTGGGTGAATGAGGTGAGGTACGGAGGCTTCTCGTGGGG	4469
Db	4195	CAGGGCCTGAAGACTAAGAGTGGGTGAATGAGGTGAGGTACGGAGGCTTCTCGTGGGG	4254
Qy	4470	GGCCGAGACCCAGGCGCTCGCCCTCGGGCCAAAGAGTTGGGGCCGCTCAGTGGAGGAGTTGTGG	4529
Db	4255	GGCCGAGACCCAGGCGCTCGCCCTCGGGCCAAAGAGTTGGGGCCGCTCAGTGGAGGAGTTGTGG	4314
Qy	4530	GGCTGTGTAGTCCCTTGCTGCGGGGGCCCTCGAACCGTGTCTGTGAAAACTCTCACAGCC	4589
Db	4315	GGCTGTGTAGTCCCTTGCTGCGGGGGCCCTCGAACCGTGTCTGTGAAAACTCTCACAGCC	4374
Qy	4590	TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAGAGGCTGG	4649
Db	4375	TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAGAGGCTGG	4434
Qy	4650	CACCTCCATGTGGCCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACTTGC	4709
Db	4435	CACCTCCATGTGGCCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACTTGC	4494
Qy	4710	CCAGGCCCGGCGCGCCACGCCCACAGCATCACACACTCAACCAACCCCTTGAACTCAC	4769
Db	4495	CCAGGCCCGGCGCGCCACGCCCACAGCATCACACACTCAACCAACCCCTTGAACTCAC	4554
Qy	4770	AAGGAGCAGCTGTGAGGCTGCACTGATGGCCCTCCTCGGTGACGCTCTCTGCTCCATC	4829
Db	4555	AAGGAGCAGCTGTGAGGCTGCACTGATGGCCCTCCTCGGTGACGCTCTCTGCTCCATC	4614
Qy	4830	TGTGTGTCTTTTGCCATGTCTTTTGTCCGGCCAGCTTCACTCTGTCTCTATTGAGGAG	4889
Db	4615	TGTGTGTCTTTTGCCATGTCTTTTGTCCGGCCAGCTTCACTCTGTCTCTATTGAGGAG	4674
Qy	4890	CGAGTCAACCGAGCCAAAGCACTGCAAGCTCATGGGGGGCTGTCCCCACCCCTCTACTGG	4949
Db	4675	CGAGTCAACCGAGCCAAAGCACTGCAAGCTCATGGGGGGCTGTCCCCACCCCTCTACTGG	4734
Qy	4950	CTTGGCAACTTCTCTGGGACATGTGTAAGTCTTGGTGGCAGCATGCTGTGGTGTCTC	5009
Db	4735	CTTGGCAACTTCTCTGGGACATGTGTAAGTCTTGGTGGCAGCATGCTGTGGTGTCTC	4794
Qy	5010	ATCTTTCTGGCCCTTCAGACAGAGGACATATGTGGCCCTTGCCAACTGCTGTCTCTCTG	5069
Db	4795	ATCTTTCTGGCCCTTCAGACAGAGGACATATGTGGCCCTTGCCAACTGCTGTCTCTCTG	4854
Qy	5070	CTGTGTCTACTACTGATGCTGGTGCATACACCGCTCATGTACCCAGCCTCTTCTTCTC	5129
Db	4855	CTGTGTCTACTACTGATGCTGGTGCATACACCGCTCATGTACCCAGCCTCTTCTTCTC	4914
Qy	5130	TTCTCGTGGCCAGCACAGCCTATGTGGTCTCACCTGCATAAACCTCTTTATTTGGCATC	5189
Db	4915	TTCTCGTGGCCAGCACAGCCTATGTGGTCTCACCTGCATAAACCTCTTTATTTGGCATC	4974
Qy	5190	AATGGAAGCATGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGAGTG	5249
Db	4975	AATGGAAGCATGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGAGTG	5034
Qy	5250	AGCCGGATCTTGAAACAGAGTCTTCTTATCTTCCCCACTTCTGTCTGGGCCGGGGGCTC	5309
Db	5035	AGCCGGATCTTGAAACAGAGTCTTCTTATCTTCCCCACTTCTGTCTGGGCCGGGGGCTC	5094
Qy	5310	ATTGACATGCTGGGAAACAGAGCCTATGGCTGATGCCCTTTGAGGCTTTGGGACAGAGCAG	5369
Db	5095	ATTGACATGCTGGGAAACAGAGCCTATGGCTGATGCCCTTTGAGGCTTTGGGACAGAGCAG	5154
Qy	5370	TTCCAGTCACTGCTGGGAGTGGTGGCAAGAACTCTTGGCCATGTGTATACAG	5429
Db	5155	TTCCAGTCACTGCTGGGAGTGGTGGCAAGAACTCTTGGCCATGTGTATACAG	5214
Qy	5430	GGGCCCCCTTCTCTTCTTCACTACTGCTGACGACCGGAAGCCAACTCTCTGCCACAG	5489

Qy	690	ACGTCACTGCTGCGGACGGAAATCCCTGGGGTTGGGACATGCGGCCAAGCCCAAGAGCCCTTG	749
Db	475	ACGTCACTGCTGCGGACGGAAATCCCTGGGGTTGGGACATGCGGCCAAGCCCAAGAGCCCTTG	534
Qy	750	CACAGCTTGTGGAGGCGCTGAGACCTGSCCCAGAGAGCTCTCGGCGCTGCGAGCGCTG	809
Db	535	CACAGCTTGTGGAGGCGCTGGGGACCTGSCCCAGAGAGCTCTCGGCGCTGCGAGCGCTG	594
Qy	810	GTGAGGCTTCGGGCACTGCTGACAGAGACCCGAGGGACCAAGCGGCCCCCTGGAGTTGCTG	869
Db	595	GTGAGGCTTCGGGCACTGCTGACAGAGACCCCGAGGGACCAAGCGGCCCCCTGGAGTTGCTG	654
Qy	870	TCAGAGGCGCTTCGAGTGTCAGGGGACCTAGACAGCAGATGGGCGCCCTCCCTCAATCGG	929
Db	655	TCAGAGGCGCTTCGAGTGTCAGGGGACCTAGACAGCAGATGGGCGCCCTCCCTCAATCGG	714
Qy	930	TACGAGGCTAGTGACCTGATGAGAGCTGTGTGGGCGAGGACAGAAATCCGCCCTGCCAGAC	989
Db	715	TACGAGGCTAGTGACCTGATGAGAGCTGTGTGGGCGAGGACAGAAATCCGCCCTGCCAGAC	774
Qy	990	AGCAGCCTGAGCCCCGCGCTGCTCGGAGCTGATTTGGAGCCCTTGGACAGCCACCCGCTGTCC	1049
Db	775	AGCAGCCTGAGCCCCGCGCTGCTCGGAGCTGATTTGGAGCCCTTGGACAGCCACCCGCTGTCC	834
Qy	1050	CGCCTGCTCTGGAGAGCGCTGAAAGCTCTGATCCCTCGGGAAGCTACTCTTTGCACCAAGAT	1109
Db	835	CGCCTGCTCTGGAGAGCGCTGAAAGCTCTGATCCCTCGGGAAGCTACTCTTTGCACCAAGAT	894
Qy	1110	ACACCTTTTACCCGGAAGCTCATCGGCCACAGGTGAACCGGACCTTCGAGGAGCTCACCCCTG	1169
Db	895	ACACCTTTTACCCGGAAGCTCATCGGCCACAGGTGAACCGGACCTTCGAGGAGCTCACCCCTG	954
Qy	1170	CTGAGGAGATGTCGCGGAGGTGTGGGAGATGTCTGGGACCCCGGATCTTTCACTTCATGAAC	1229
Db	955	CTGAGGAGATGTCGCGGAGGTGTGGGAGATGTCTGGGACCCCGGATCTTTCACTTCATGAAC	1014
Qy	1230	GACAGTTCCAAATGTGGCCATGCTGCAGCGGCTCTGTCAGATGCAAGATGAAGGAAGAAGG	1289
Db	1015	GACAGTTCCAAATGTGGCCATGCTGCAGCGGCTCTGTCAGATGCAAGATGAAGGAAGAAGG	1074
Qy	1290	CAGCCAGACCTGGAGCGGGACCAATGGAGGCCCTCGATCTTTCTGGAACCTCTGG	1349
Db	1075	CAGCCAGACCTGGAGCGGGACCAATGGAGGCCCTCGATCTTTCTGGAACCTCTGG	1134
Qy	1350	AGCGGTGGCTACAGCTGGCAGGACGCACACGCTGATGTGGGGACCTTGGTGGGACCGCTG	1409
Db	1135	AGCGGTGGCTACAGCTGGCAGGACGCACACGCTGATGTGGGGACCTTGGTGGGACCGCTG	1194
Qy	1410	GGCCGAGTGACGAGTAGTCTGTCTTTGGACAAAGCTGAGGCGGCAACCTCAGAGGACGCC	1469
Db	1195	GGCCGAGTGACGAGTAGTCTGTCTTTGGACAAAGCTGAGGCGGCAACCTCAGAGGACGCC	1254
Qy	1470	CTGTGTGTCGGGGCCCTGCACCTCTCGCGGAACATCGATCTTGGGCGCGGCTGCTTC	1529
Db	1255	CTGTGTGTCGGGGCCCTGCACCTCTCGCGGAACATCGATCTTGGGCGCGGCTGCTTC	1314
Qy	1530	TTGGGACCTGAGGACTCTTCAGACCCCAAGAGACCCCAACCTTCGGGCGCGGCTGCTTC	1589
Db	1315	TTGGGACCTGAGGACTCTTCAGACCCCAAGAGACCCCAACCTTCGGGCGCGGCTGCTTC	1374
Qy	1590	CAGCTGCGCATCAAAATCCGCATGGAATTTGAGTGCTGTCAGAGGACCAATAAGATCAGG	1649
Db	1375	CAGCTGCGCATCAAAATCCGCATGGAATTTGAGTGCTGTCAGAGGACCAATAAGATCAGG	1434
Qy	1650	GACAGGTTTTGGGACCTTGGCCGAGCGCGGACCCCTGACCCGACCTGCGCTACGTTGG	1709
Db	1435	GACAGGTTTTGGGACCTTGGCCGAGCGCGGACCCCTGACCCGACCTGCGCTACGTTGG	1494
Qy	1710	GGCGGCTTCGTGTACCTGCAAGACCTGCTGGAGCGTGCAGCCGCTCCGCTCAGCGGC	1769
Db	1495	GGCGGCTTCGTGTACCTGCAAGACCTGCTGGAGCGTGCAGCCGCTCCGCTCAGCGGC	1554
Qy	1770	GCACACCCCGGGCGGCGCTCTACTGACAGCAGATGCCCTTATCCGCTGCTATGTGGACAC	1829

Db	1555	GCCAACCCCGGGCGCGCCTTACTCTGACAGATGCCCTATCTCGTGTCTATGTGGACGAC	1614
Qy	1830	GTGTTCTGTGGTGTGCTGAGCCGCTGCTGCGCTCTTCTCTGACGCTGGCTGGATCTAC	1889
Db	1615	GTGTTCTGTGGTGTGCTGAGCCGCTGCTGCGCTCTTCTCTGACGCTGGCTGGATCTAC	1674
Qy	1890	TCCGTGACACTGACAGTGAAGCGCGTGGTTCGGGAGAGAGACGCGGCTGCGGGACACC	1949
Db	1675	TCCGTGACACTGACAGTGAAGCGCGTGGTTCGGGAGAGAGACGCGGCTGCGGGACACC	1734
Qy	1950	ATGCGCGCATGGGGCTCAGCGCGCGGGTGGTCTGGCTAGGCTGGTTCCTCAGCTGCGCTC	2009
Db	1735	ATGCGCGCATGGGGCTCAGCGCGCGGGTGGTCTGGCTAGGCTGGTTCCTCAGCTGCGCTC	1794
Qy	2010	GGGCGCTTCTGTGCTCAGCGCGCGCTGCTGGTTCGTGGTCTCAAGCTGGGGGACATCCTC	2069
Db	1795	GGGCGCTTCTGTGCTCAGCGCGCGCTGCTGGTTCGTGGTCTCAAGCTGGGGGACATCCTC	1854
Qy	2070	CCCTACAGCCAACCGCGCGTGGTTCCTCTGTTCTTGGCAGCCTTTCGCGGTGGCCACGGTG	2129
Db	1855	CCCTACAGCCAACCGCGCGTGGTTCCTCTGTTCTTGGCAGCCTTTCGCGGTGGCCACGGTG	1914
Qy	2130	ACCAGAGCTTCTGTCTCAGGGCTTCTTCTCCGGCGCCAACTGGCTGGGCTGGCGCTGGCGC	2189
Db	1915	ACCAGAGCTTCTGTCTCAGGGCTTCTTCTCCGGCGCCAACTGGCTGGGCTGGCGCTGGCGC	1974
Qy	2190	GGCTGCGCTACTTCTCCCTTACTCCCTTACTGCTGCTGTGTGGCTTCGCGGGACCGG	2249
Db	1975	GGCTGCGCTACTTCTCCCTTACTGCTGCTGTGTGGCTTCGCGGGACCGG	2034
Qy	2250	CTGCCCGCGGGTGCGCGCTGGCGGAGCCTGTGTGCGCGCTTCGGCTTCGCGCTCGGC	2309
Db	2035	CTGCCCGCGGGTGCGCGCTGGCGGAGCCTGTGTGCGCGCTTCGGCTTCGCGCTCGGC	2094
Qy	2310	TGCAGAGCCTGGCTCTGTGTGAGAGACAGGGCGAGGGCGCGAGTGGGCAACAGTGGGC	2369
Db	2095	TGCAGAGCCTGGCTCTGTGTGAGAGACAGGGCGAGGGCGCGAGTGGGCAACAGTGGGC	2154
Qy	2370	ACCGGCGCTACGCAGACGCTTTCAGCCTGGCGCCAGCTCTCTGGCCTTCTGCTGTGGAC	2429
Db	2155	ACCGGCGCTACGCAGACGCTTTCAGCCTGGCGCCAGCTCTCTGGCCTTCTGCTGTGGAC	2214
Qy	2430	GGGCGCTCTACGGCCTTCGCCACCTGGTACTTGGAAAGCTGTGTGCCAGGCCAGTACGGG	2489
Db	2215	GGGCGCTCTACGGCCTTCGCCACCTGGTACTTGGAAAGCTGTGTGCCAGGCCAGTACGGG	2274
Qy	2490	ATCCTGAAACATGGAAATTTTCTTTTCGAGAGAGCTACTGTGGTCGGACCTCGCGCCCCC	2549
Db	2275	ATCCTGAAACATGGAAATTTTCTTTTCGAGAGAGCTACTGTGGTCGGACCTCGCGCCCCC	2334
Qy	2550	AAGAGTCCAGCCCTTCCGCCACCCCGCTGGACCCCAAAGTGTGTGTAGAAAGGACACCG	2609
Db	2335	AAGAGTCCAGCCCTTCCGCCACCCCGCTGGACCCCAAAGTGTGTGTAGAAAGGACACCG	2394
Qy	2610	CCGCGCTGAGTCTCTGGCGTATCCGTTTCGAGCCTGGAGAGCGCTTTCCTGGAAGCCCG	2669
Db	2395	CCGCGCTGAGTCTCTGGCGTCTCCGTTTCGAGCCTGGAGAGCGCTTTCCTGGAAGCCCG	2454
Qy	2670	CAGCAGCCCTCGGGGGCTCAGCTGTGAATTCTTTCGAGAGAGCTACTGTGGTCGGACCTCGCGCCCC	2729
Db	2455	CAGCAGCCCTCGGGGGCTCAGCTGTGAATTCTTTCGAGAGAGCTACTGTGGTCGGACCTCGCGCCCC	2514
Qy	2730	GGCCACAACGGGCGCGCAAGACCAACCCCTGTCCATCTTGTAGTGGCTCTTCCCAACC	2789
Db	2515	GGCCACAACGGGCGCGCAAGACCAACCCCTGTCCATCTTGTAGTGGCTCTTCCCAACC	2574
Qy	2790	AGTGGTGGCTCTGCTTTCATCTCTGGGCGACAGAGTCCGCTCCAGCATGGCGGCATCCCGG	2849
Db	2575	AGTGGTGGCTCTGCTTTCATCTCTGGGCGACAGAGTCCGCTCCAGCATGGCGGCATCCCGG	2634
Qy	2850	CCCCACTGGGCGTGTGCTCTAGTACAACGCTGCTGTTTGACATGCTGACCGGTGGACGAG	2909

QY 5070 CTGTTGCTACTGTTATGGCTGTCGATCACACCGCTCATGTACCCAGCCTCTTCTTC 5129
Db 4855 CTGTTGCTACTACTGTTATGGCTGTCGATCACACCGCTCATGTACCCAGCCTCTTCTTC 4914
QY 5130 TTCTCCGTCGCCAGCACAGCCTATGTGTCCTCACTGCATAAACCTCTTATTGGCATC 5189
Db 4915 TTCTCCGTCGCCAGCACAGCCTATGTGTCCTCACTGCATAAACCTCTTATTGGCATC 4974
QY 5190 AATGGAAGCATGGCCACCTTTGTGCTTGAGCTCTTCTGTATCAAGAGCTGCAGAGGTG 5249
Db 4975 AATGGAAGCATGGCCACCTTTGTGCTTGAGCTCTTCTGTATCAAGAGCTGCAGAGGTG 5034
QY 5250 AGCCGGATCTTGAACAGAGCTTCTTATCTTCCCCCACTTCTGTGTCGGCCGGGGCTC 5309
Db 5035 AGCCGGATCTTGAACAGAGCTTCTTATCTTCCCCCACTTCTGTGTCGGCCGGGGCTC 5094
QY 5310 ATTGACATGGTCGGAAACAGGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAG 5369
Db 5095 ATTGACATGGTCGGAAACAGGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAG 5154
QY 5370 TTCCAGTCAACCCCTGCGCTGGAGGTGTCGCAAGAACCTCTTGGCCCATGCTGATACAG 5429
Db 5155 TTCCAGTCAACCCCTGCGCTGGAGGTGTCGCAAGAACCTCTTGGCCCATGCTGATACAG 5214
QY 5430 GGGCCCTCTCTCTTCTTCACTACTGTGACAGCCGAAGCAACTCTTGCACAG 5489
Db 5215 GGGCCCTCTCTCTTCTTCACTACTGTGACAGCCGAAGCAACTCTTGCACAG 5274
QY 5490 CCCAGGTGAGTCTCTGCACTCTGCGAGGAGGAGGAGGATGTAGCCGCTGAACGG 5549
Db 5275 CCCAGGTGAGTCTCTGCACTCTGCGAGGAGGAGGAGGATGTAGCCGCTGAACGG 5334
QY 5550 GAGCGGTGTCCTAAGAGGACCAACAGGGGATGTGTTGGTCTGAGGAACTTGACCAAG 5609
Db 5335 GAGCGGTGTCCTAAGAGGACCAACAGGGGATGTGTTGGTCTGAGGAACTTGACCAAG 5394
QY 5610 GTATACCGTGGCAGAGGATGCCAGCTTTGACCGCTTTGCTGGGATTTCCCTCTGT 5669
Db 5395 GTATACCGTGGCAGAGGATGCCAGCTTTGACCGCTTTGCTGGGATTTCCCTCTGT 5454
QY 5670 GAGTGTGTTGGCTGCTGGTGTGTAATGAGCAGGGAAGAGTCCAGTTTCGATGCTG 5729
Db 5455 GAGTGTGTTGGCTGCTGGTGTGTAATGAGCAGGGAAGAGTCCAGTTTCGATGCTG 5514
QY 5730 ACGGGGACATATGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCCACAGGTTGCCCGG 5789
Db 5515 ACGGGGACATATGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCCACAGGTTGCCCGG 5574
QY 5790 GAAACCACTGCTGCGCACCTCAGCATGGGATATGCCCCCTCAATCCGATGCCATCTTTGAG 5849
Db 5575 GAAACCACTGCTGCGCACCTCAGCATGGGATATGCCCCCTCAATCCGATGCCATCTTTGAG 5634
QY 5850 CTGCTGACGGGCGGGAGACACTGAGAGTCTTGGCGGCTGCGCGGTGCCGAGGCC 5909
Db 5635 CTGCTGACGGGCGGGAGACACTGAGAGTCTTGGCGGCTGCGCGGTGCCGAGGCC 5694
QY 5910 CAGTTGCCAGACCGCTGCTCAGGCTGCGGCTGCTGGACTCTCATGTAGCCAGAC 5969
Db 5695 CAGTTGCCAGACCGCTGCTCAGGCTGCGGCTGCTGGACTCTCATGTAGCCAGAC 5754
QY 5970 CGGCTCGAGGACCTTACAGCGGGGGAACAAACGCAAGCTGGGACCGGCCCTGCGCGTG 6029
Db 5755 CGGCTCGAGGACCTTACAGCGGGGGAACAAACGCAAGCTGGGACCGGCCCTGCGCGTG 5814
QY 6030 GTTGGGACCCAGCGCTGTTCTGGAACAGCGCCGACCAAGGATGGACCCAGCGCG 6089
Db 5815 GTTGGGACCCAGCGCTGTTCTGGAACAGCGCCGACCAAGGATGGACCCAGCGCG 5874
QY 6090 CGGCGCTCTCTTGGAAACAGCTTTGCGCTGCTGGGAGGCGCTTCACTGATGCTC 6149
Db 5875 CGGCGCTCTCTTGGAAACAGCTTTGCGCTGCTGGGAGGCGCTTCACTGATGCTC 5934
QY 6150 ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGCTGAAT 6209

Db 5935 ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCCCTAGCCATCATGCTGAAT 5994
QY 6210 GGGCGGTTCCGCTCGCTGGGAGCCGCAACATCTCAAGGCGAGATTCCGCGCGGTCAC 6269
Db 5995 GGGCGGTTCCGCTCGCTGGGAGCCGCAACATCTCAAGGCGAGATTCCGCGCGGTCAC 6054
QY 6270 AACTGACCTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCTGTCGCGCC 6329
Db 6055 AACTGACCTCTGCGGCTGCGGCTGCGGCTGCGGCTTCTGTCGCGCC 6114
QY 6330 GAGTTCCCTGGTTCGAGCTGCGGAGGACATGAGGCGCCCTGCGCTTCCAGTTCGG 6389
Db 6115 GAGTTCCCTGGTTCGAGCTGCGGAGGACATGAGGCTGCGCTTCCAGTTCGG 6174
QY 6390 CCGGAGGCGCTGCGCCCTGCGGCTGCGGCTGCGGCTGCGGCTTCTGTCGCGCC 6449
Db 6175 CCGGAGGCGCTGCGCCCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGAGT 6234
QY 6450 CACGCGTGGAGGACTTTTTCGCTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTC 6509
Db 6235 CACGCGTGGAGGACTTTTTCGCTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTC 6294
QY 6510 TCCAAGACACAGGGAAGGAGGACACCGAAGACGAAAGGAGGAGTGGAGTG 6569
Db 6295 TCCAAGACACAGGGAAGGAGGAGGACACCGAAGACGAAAGGAGGAGTGGAGTG 6354
QY 6570 GACCCCGCGCAGGCTGCGACACCCAAACGCTGAGCCAGTTCTCGATGACCTTAGC 6629
Db 6355 GACCCCGCGCAGGCTGCGACACCCAAACGCTGAGCCAGTTCTCGATGACCTTAGC 6414
QY 6630 ACTGCCGAGACTGTGCTC 6647
Db 6415 ACTGCCGAGACTGTGCTC 6432

RESULT 12

US-10-775-920-6
; Sequence 6, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 6174
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-6

Query Match 89.3%; Score 6073.6; DB 17; Length 6174;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6082; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 702 CGCACGGAATCCCTGGGGTTGGCACTGGCCCAAGCCAGGAGCCTCTGCGCTGCGCAGCCTGGTGGAGCTTCGG 761
Db 79 CCCAGGAATCCCTGGGGTTGGCACTGGCCCAAGCCAGGAGCCTCTGCGCTGCGCAGCCTGGTGGAGCTTCGG 138
QY 762 GAGCGCGCTGAGGACCTGGCCCGAGGAGCTCTGCGCTGCGCAGCCTGGTGGAGCTTCGG 821
Db 139 GAGCGCGCTGAGGACCTGGCCCGAGGAGCTCTGCGCTGCGCAGCCTGGTGGAGCTTCGG 198
QY 822 GCATGCTGACAGACCCCGAGGAGCCAGCGGCCCTCTGGAGTTGCTGTACAGGCCCTC 881
Db 199 GCATGCTGACAGACCCCGAGGAGCCAGCGGCCCTCTGGAGTTGCTGTACAGGCCCTC 258

QY 882 TGCAAGTGTACAGGAGCCTAGCAGCACAGTGGGCCCCCTCCCTCAACTGGTACGAGGCTAGT 941
DB 259 TGCAAGTGTACAGGAGCCTAGCAGCACAGTGGGCCCCCTCCCTCAACTGGTACGAGGCTAGT 318
QY 942 GACCTGATGAGCTGTGTGGGCGAGGAGCCAGAAATCCGCCCCCTGCCAGACAGCAGCCTGAGC 1001
DB 319 GACCTGATGAGCTGTGTGGGCGAGGAGCCAGAAATCCGCCCCCTGCCAGACAGCAGCCTGAGC 378
QY 1002 CCGCCCTGCTCGAGCTGATGTGAGCCCTGGACAGCCACCCGCTGTCCCGCCTGCTCTGG 1061
DB 379 CCGCCCTGCTCGAGCTGATGTGAGCCCTGGACAGCCACCCGCTGTCCCGCCTGCTCTGG 438
QY 1062 AGACGCTGAAGCTCTGATCTCTGGGAAGCTACTTTTTCACACAGATACACCTTTTACC 1121
DB 439 AGACGCTGAAGCTCTGATCTCTGGGAAGCTACTTTTTCACACAGATACACCTTTTACC 498
QY 1122 CGAAAGCTCATGCCCCCAGGTGAACCGGACCTTCGAGAGCTCACCTCTGCTGAGGGATGTC 1181
DB 499 CGAAAGCTCATGCCCCCAGGTCAACCGGACCTTCGAGAGCTCACCTCTGCTGAGGGATGTC 558
QY 1182 CGGAGGTGTGGAGATGCTGGGAGCCCCCGGATCTTACCTTCATGAACGACAGTTCGAAT 1241
DB 559 CGGAGGTGTGGAGATGCTGGGAGCCCCCGGATCTTACCTTCATGAACGACAGTTCGAAT 618
QY 1242 GTGGCCATGCTGCAGCGGCTCCTGCGAGATCGAGATGAAGGAAGAGGAGCCAGACCT 1301
DB 619 GTGGCCATGCTGCAGCGGCTCCTGCGAGATCGAGATGAAGGAAGAGGAGCCAGACCT 678
QY 1302 GGAGGCGGGACACATGAGGAGCCCTGCGATCTTCTGACCCCTGGGAGCGGTGGCTAC 1361
DB 679 GGAGGCGGGACACATGAGGAGCCCTGCGATCTTCTGACCCCTGGGAGCGGTGGCTAC 738
QY 1362 AGCTGGCAGACGACACGCTGATGTGGGCACTGGTGGGCAAGCTGGGCGAGTGACG 1421
DB 739 AGCTGGCAGACGACACGCTGATGTGGGCACTGGTGGGCAAGCTGGGCGAGTGACG 798
QY 1422 GAGTGCCTGTCTTGGACAGCTGGAGGCGGACCCCTCAGAGCAGCCCTGCTGTGCGG 1481
DB 799 GAGTGCCTGTCTTGGACAGCTGGAGGCGGACCCCTCAGAGCAGCCCTGCTGTGCGG 858
QY 1482 GCCTTGCAACTGTCTGCGGGAACATCGATTTCTGGGCGGCGCTCGTCTTCTTGGACCTGAG 1541
DB 859 GCCTTGCAACTGTCTGCGGGAACATCGATTTCTGGGCGGCGCTCGTCTTCTTGGACCTGAG 918
QY 1542 GACTTTTCAGACCCCAAGAGCACCACCCAGACCTGGGCCCCCGGCAAGTGGGATC 1601
DB 919 GACTTTTCAGACCCCAAGAGCACCACCCAGACCTGGGCCCCCGGCAAGTGGGATC 978
QY 1602 ABAATCGCATGACATTGACGTGTCAGAGGACCAATAAGATCAGGACAGTGTG 1661
DB 979 ABAATCGCATGACATTGACGTGTCAGAGGACCAATAAGATCAGGACAGTGTG 1038
QY 1662 GACCTTGGCCAGCGCGGACCCCTGACCGACTGCGCTACGCTGAGTGGGCGGCTTCGTG 1721
DB 1039 GACCTTGGCCAGCGCGGACCCCTGACCGACTGCGCTACGCTGAGTGGGCGGCTTCGTG 1098
QY 1722 TACTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCAAGCGCGCAACCCCGG 1781
DB 1099 TACTGCAAGACCTGTGTGAGCGTGCAGCGCTCCGCGTGTCAAGCGCGCAACCCCGG 1158
QY 1782 GCGGCGCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGACGACGCTGCTGCGT 1841
DB 1159 GCGGCGCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGACGACGCTGCTGCGT 1218
QY 1842 GTGCTGAGCGGCTCGCTGCGGCTCTTCTGACGCTGGCCTGGATCTACTCCGTGACACTG 1901
DB 1219 GTGCTGAGCGGCTCGCTGCGGCTCTTCTGACGCTGGCCTGGATCTACTCCGTGACACTG 1278
QY 1902 ACAGTGAAGCCGCTGTGTGGGAGAAAGGAGACCGCGCTGCGGACACCAATGCGCGCATG 1961
DB 1279 ACAGTGAAGCCGCTGTGTGGGAGAAAGGAGACCGCGCTGCGGACACCAATGCGCGCATG 1338

QY 1962 GGGCTCAGCGCGCGGTGCTCTGGCTAGGCTGGTTCTCAGCTGCTCAGGCGCTTCCTG 2021
DB 1339 GGGCTCAGCGCGCGGTGCTCTGGCTAGGCTGGTTCTCAGCTGCTCAGGCGCTTCCTG 1398
QY 2022 CTGAGCGCGCGGTGCTCTGGTTCAGAGCTGGGGAACATCTCTCCCTTACAGCCAC 2081
DB 1399 CTGAGCGCGCGGTGCTCTGGTTCAGAGCTGGGGAACATCTCTCCCTTACAGCCAC 1458
QY 2082 CCGGCGGTGGTCTTCTGCTTCTGGCAGCCTTTCGCGTGGCCACAGTGGACCCAGAGCTTC 2141
DB 1459 CCGGCGGTGGTCTTCTGCTTCTGGCAGCCTTTCGCGTGGCCACAGTGGACCCAGAGCTTC 1518
QY 2142 CTGCTCAGCGCGCTTCTTCTCCGCGCAACCTGGCTGCGGCTTCGCGCGGCTTCGGCTAC 2201
DB 1519 CTGCTCAGCGCGCTTCTTCTCCGCGCAACCTGGCTGCGGCTTCGCGCGGCTTCGGCTAC 1578
QY 2202 TTTCTCCTCTACCTGCCCCCTAGCTGTGTGTGGCTTTGGCGGACCGGCTGCCCGGGT 2261
DB 1579 TTTCTCCTCTACCTGCCCCCTAGCTGTGTGTGGCTTTGGCGGACCGGCTGCCCGGGT 1638
QY 2262 GCGCGGTGGCGCGAGCCTGCTCGCCGCTGGCCTTCGGCTTCGGCTTCGCGAGAGCCTG 2321
DB 1639 GCGCGGTGGCGCGAGCCTGCTCGCCGCTGGCCTTCGGCTTCGGCTTCGCGAGAGCCTG 1698
QY 2322 GCTCTCTGGAGAGCAGCGCGAGGCGCGAGTGGCAACACGTGGGCAACCCGCGCTACG 2381
DB 1699 GCTCTCTGGAGAGCAGCGCGAGGCGCGAGTGGCAACACGTGGGCAACCCGCGCTACG 1758
QY 2382 GCAGAGCTTTTACGCTGCCCCAGGFTCTCTGGCTTTCTGCTGCAACCGCGGCTCTAC 2441
DB 1759 GCAGAGCTTTTACGCTGCCCCAGGFTCTCTGGCTTTCTGCTGCAACCGCGGCTCTAC 1818
QY 2442 GGGCTCGCACCTGCTACCTGGAAGCTGTGCCCCAGGCGAGTACGGGATCCCTGAACCA 2501
DB 1819 GGGCTCGCACCTGCTACCTGGAAGCTGTGCCCCAGGCGAGTACGGGATCCCTGAACCA 1878
QY 2502 TGGAAATTTTCTTTTCGAGAGGACTACTGCTGGGACCTTCGGCCCCCAGAGTCCAGCC 2561
DB 1879 TGGAAATTTTCTTTTCGAGAGGACTACTGCTGGGACCTTCGGCCCCCAGAGTCCAGCC 1938
QY 2562 CTTTGCCCCACCCCGCTGGACCCAAAGTGTCTGTAGAGAGGACACCGCCCCGCGCTGAGT 2621
DB 1939 CTTTGCCCCACCCCGCTGGACCCAAAGTGTCTGTAGAGAGGACACCGCCCCGCGCTGAGT 1998
QY 2622 CTTGGGCTATCCGTTTCGAGCCTTGGAGAGCGCTTCTGGAAGCGCGCGAGCCAGCCCTG 2681
DB 1999 CTTGGGCTATCCGTTTCGAGCCTTGGAGAGCGCTTCTGGAAGCGCGCGAGCCAGCCCTG 2058
QY 2682 CCGGCGCTCAGCTGACCTTACACAGGCGCACATCACCGCCTTCCTGGGCGCAACAGGG 2741
DB 2059 CCGGCGCTCAGCTGACCTTACACAGGCGCACATCACCGCCTTCCTGGGCGCAACAGGG 2118
QY 2742 GCGGCAAGACCAACCACTGCTTCATCTTTGAGTGGCTCTTCCCAACCCAGTGGTCTCT 2801
DB 2119 GCGGCAAGACCAACCACTGCTTCATCTTTGAGTGGCTCTTCTCCCAACCCAGTGGTCTCT 2178
QY 2802 GCTTCATCTGGGCGACAGCTCGCTCCAGATGCGCGCATCCGCGCCCCACCTGGGC 2861
DB 2179 GCTTCATCTGGGCGACAGCTCGCTCCAGATGCGCGCATCCGCGCCCCACCTGGGC 2238
QY 2862 GTCTGTCTCAGTACAAACGCTGTGTTGATGCTGACCTGCGTGGACAGACGCTGCTGTT 2921
DB 2239 GTCTGTCTCAGTACAAACGCTGTGTTGATGCTGACCTGCGTGGACAGACGCTGCTGTT 2298
QY 2922 TATGGCGGCTGAAGGCTGTGAGTCCGCTGTAGTGGGCCCCCGAGCAGGACCGTCTGCTG 2981
DB 2299 TATGGCGGCTGAAGGCTGTGAGTCCGCTGTAGTGGGCCCCCGAGCAGGACCGTCTGCTG 2358
QY 2982 CAGGATGTGGGCTGTCTCCAGCAGAGTGTGACACTTCGCCACTCTCTGTGGGATG 3041
DB 2359 CAGGATGTGGGCTGTCTCCAGCAGAGTGTGACACTTCGCCACTCTCTGTGGGATG 2418
QY 3042 CAACGGAAGCTGTCTGCTGGCCATTGCTTTGTGGGCGGCTCCCAAGTGTGTTATCTCTGGAC 3101

Db 2419 CAAAGGAGCTGTCGGTGCCATTTGCTTTGGGGGGCTCCAAAGTTGTTATCTCGAC 2478
QY 3102 GAGCCTACGGCTGGCGTGATCCTGCTTCCCGCGGTAATTTGGGAGCTGCTGTCAAA 3161
Db 2479 GAGCCTACGGCTGGCGTGATCCTGCTTCCCGCGGTAATTTGGGAGCTGCTGTCAAA 2538
QY 3162 TACCAAGAGGTGGACCGTGAATCTCTCCACCAACACTGGATGAGGAGAGTGTGTG 3221
Db 2539 TACCAAGAGGTGGACCGTGAATCTCTCCACCAACACTGGATGAGGAGAGTGTGTG 2598
QY 3222 GAGACCGGTGGCTGGTGGAGGAGTGGCGGCTTGTGTGCTGTGGCTCCCACTCTTC 3281
Db 2599 GAGACCGGTGGCTGGTGGAGGAGTGGCGGCTTGTGTGCTGTGGCTCCCACTCTTC 2658
QY 3282 GTGCGCGCTCACCTGGGCTCCGGCTACTACCTGAGCTGGTGAAGGCGCGCTGCCCTG 3341
Db 2659 GTGCGCGCTCACCTGGGCTCCGGCTACTACCTGAGCTGGTGAAGGCGCGCTGCCCTG 2718
QY 3342 ACCACCAATGAGAAGCTGACACTGACATGAGAGGAGTGTGGACACAGGAGGAAAG 3401
Db 2719 ACCACCAATGAGAAGCTGACACTGACATGAGAGGAGTGTGGACACAGGAGGAAAG 2778
QY 3402 AAGATGCGACGACGAGGAGAGTGGCGACTCTCAGCTGCTGGCCCTGGTACAGCAC 3461
Db 2779 AAGATGCGACGAGGAGGAGAGTGGCGACTCTCAGCTGCTGGCCCTGGTACAGCAC 2838
QY 3462 TGGGTGCCCGGGGACGGCTGGTGAGAGCTGGCACAGAGCTGGTGTGCTGCGCC 3521
Db 2839 TGGGTGCCCGGGGACGGCTGGTGAGAGCTGGCACAGAGCTGGTGTGCTGCGCC 2898
QY 3522 TACACGGGTGCCATGACGGAGCTTCGCCACACTCTTCCGAGAGCTAGACACGGGCTG 3581
Db 2899 TACACGGGTGCCATGACGGAGCTTCGCCACACTCTTCCGAGAGCTAGACACGGGCTG 2958
QY 3582 GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAGGAGATCTTCTG 3641
Db 2959 GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAGGAGATCTTCTG 3018
QY 3642 AAGTGTGGAGAGTGTGCTGCGACACAGATATGAGAGTGGCAGCTGGGGGAGCAC 3701
Db 3019 AAGTGTGGAGAGTGTGCTGCGACACAGATATGAGAGTGGCAGCTGGGGGAGCAC 3078
QY 3702 CTATGACAGGATTTGCTGGCTAGACGTAAAGCTTGGGCTCAAGATGCCGCCACAGGAG 3761
Db 3079 CTATGACAGGATTTGCTGGCTAGACGTAAAGCTTGGGCTCAAGATGCCGCCACAGGAG 3138
QY 3762 ACAGCGCTGAGAACGGGGAACAGCTGGGTACGCCCCAGAGACTGACAGGCTCTGGG 3821
Db 3139 ACAGCGCTGAGAACGGGGAACAGCTGGGTACGCCCCAGAGACTGACAGGCTCTGGG 3198
QY 3822 CCAGACCGCTGGGCGGGTACAGGCTGGGCACTGACCGCCAGCAGCTCCAGGCCCTG 3881
Db 3199 CCAGACCGCTGGGCGGGTACAGGCTGGGCACTGACCGCCAGCAGCTCCAGGCCCTG 3258
QY 3882 CTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGCTGTTGCGCCAGATGCTG 3941
Db 3259 CTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGCTGTTGCGCCAGATGCTG 3318
QY 3942 CTGCTGCCCTTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATGTCGCTCTTTTGGG 4001
Db 3319 CTGCTGCCCTTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATGTCGCTCTTTTGGG 3378
QY 4002 CACTACCGGCTCTGCGGCTCAGTCCCAACATGATGAGTGTCTCTTTCAGT 4061
Db 3379 CACTACCGGCTCTGCGGCTCAGTCCCAACATGATGAGTGTCTCTTTCAGT 3438
QY 4062 GAGGACGCCAGAGGAGCCCTGAGCTGCGGCTGCTCGAGGCGCTGCTCGAGGAGCA 4121
Db 3439 GAGGACGCCAGAGGAGCCCTGAGCTGCGGCTGCTCGAGGCGCTGCTCGAGGAGCA 3498
QY 4122 GGACTGAGGAGGAGCCCGCAGTACAGTATCCCAAGTTCTCGGACACAGAGTTCTCT 4181

Db 3499 GGACTGAGGAGGAGCCCGCAGTACAGTATAGTCTCCACAGGTTCTCGGACACAGAGTTCT 3558
QY 4182 GCTGAAGTGGCAAGAGTCTTGGCAGTGGCAACTGGAGCTCTCCATCCCGAGCC 4241
Db 3559 GCTGAAGTGGCAAGAGTCTTGGCAGTGGCAACTGGAGCTCTCCATCCCGAGCC 3618
QY 4242 TGCCAGTGTAGCAGGAGCCCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCTG 4301
Db 3619 TGCCAGTGTAGCAGGAGCCCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCTG 3678
QY 4302 GGTCCCGCTCGCGCCAGGAGTGAACCGGCTCTGGGAGTGGTTCAAGACTGACAGG 4361
Db 3679 GGTCCCGCTCGCGCCAGGAGTGAACCGGCTCTGGGAGTGGTTCAAGACTGACAGG 3738
QY 4362 CGGAACCTGTCTGACTTCTGCTCAAGACTTACCGCGCTGGTGGCGCAGGCGCTGAAG 4421
Db 3739 CGGAACCTGTCTGACTTCTGCTCAAGACTTACCGCGCTGGTGGCGCAGGCGCTGAAG 3798
QY 4422 ACTAAGAGTGGGTGAATGAGGTACGGTACGGAGCTTCTCGTGGGGGGCGGAGCCCA 4481
Db 3799 ACTAAGAGTGGGTGAATGAGGTACGGTACGGAGCTTCTCGTGGGGGGCGGAGCCCA 3858
QY 4482 GGCCTGCCCTCGGGCCAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGT 4541
Db 3859 GGCCTGCCCTCGGGCCAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGT 3918
QY 4542 CCCCTGCTCGGGGGCCCTCGACCGTCTGCTGAAACCTCAGAGCTGGGCTCAGAGC 4601
Db 3919 CCCCTGCTCGGGGGCCCTCGACCGTCTGCTGAAACCTCAGAGCTGGGCTCAGAGC 3978
QY 4602 CTGATGCTCAGGAGCAGTCTCAAGATCTGGTTCAACAAAGGCTGGCACTCCATGGTG 4661
Db 3979 CTGAGCGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGGCACTCCATGGTG 4038
QY 4662 GCGTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCGCCCGGAGCCCGGCG 4721
Db 4039 GCGTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCGCCCGGAGCCCGGCG 4098
QY 4722 CGCCAGCCACAGCATCACCACTCAACCACTCAACCACTCAACCACTCAACCACTCAACCA 4781
Db 4099 CGCCAGCCACAGCATCACCACTCAACCACTCAACCACTCAACCACTCAACCACTCAACCA 4158
QY 4782 TCTGAGGTGCACTGATGGGCTCTCGGTGAGAGCTCTCGTCTCATCTGTGTGTCTTT 4841
Db 4159 TTTGAGGCTGATGATGGGCTCTCGGTGAGAGCTCTCGTCTCATCTGTGTGTCTTT 4218
QY 4842 GCGATGCTTTGTTCGGGCGAGCTTCACTTGTCTCTCATTTGAGGAGGAGTCAACCGA 4901
Db 4219 GCGATGCTTTGTTCGGGCGAGCTTCACTTGTCTCTCATTTGAGGAGGAGTCAACCGA 4278
QY 4902 GCGAAGCAGCTGAGCTCATGGGGGGCTGTCGCCCACTCTACTGCTTGGCACTTT 4961
Db 4279 GCGAAGCAGCTGAGCTCATGGGGGGCTGTCGCCCACTCTACTGCTTGGCACTTT 4338
QY 4962 CTCTGGACATGTTGTAATCTACTTGGTGCAGCATGCTGCTGCTCATCTTTTGCGC 5021
Db 4339 CTCTGGACATGTTGTAATCTACTTGGTGCAGCATGCTGCTGCTCATCTTTTGCGC 4398
QY 5022 TTCCAGAGGAGGAGTATGGGCGCTGCGCACTGCTGCTCTCTGCTGTGCTACTA 5081
Db 4399 TTCCAGAGGAGGAGTATGGGCGCTGCGCACTGCTGCTCTCTGCTGTGCTACTA 4458
QY 5082 CTGTATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5141
Db 4459 CTGTATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4518
QY 5142 AGCAGCCTATGTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5201
Db 4519 AGCAGCCTATGTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4578
QY 5202 GCGACCTTGTGCTGAGCTCTTCTCTGATCAGAGCTGAGGAGGAGGAGGAGGAGGAGG 5261
Db 4579 GCGACCTTGTGCTGAGCTCTTCTCTGATCAGAGCTGAGGAGGAGGAGGAGGAGGAGG 4638

QY 5262 AAACAGGTCTTCTTATCTTCCCACTTCTGCTTGGCGGGGCTCATTTGACATGGTG 5321
DB 4639 AAACAGGTCTTCTTATCTTCCCACTTCTGCTTGGCGGGGCTTATTTGACATGGTG 4698
QY 5322 CGGAACAGGCCATGGCTGATGCTTTAGCGCTTGGGAGACAGGCAAGTTCCAGTCACCC 5381
DB 4699 CGGAACAGGCCATGGCTGATGCTTTAGCGCTTGGGAGACAGGCAAGTTCCAGTCACCC 4758
QY 5382 CTGGCTTGGAGGTGGTGGCAAGAACCTCTTGGCCATGCTGATACAGGGGCCCTCTTC 5441
DB 4759 CTGGCTTGGAGGTGGTGGCAAGAACCTCTTGGCCATGCTGATACAGGGGCCCTCTTC 4818
QY 5442 CTTCTCTTCACTACTGCTGACACCGAAGCCAACTCTCTGCCACACAGCCACGGTGAGG 5501
DB 4819 CTTCTCTTCACTACTGCTGACACCGAAGCCAACTCTCTGCCACACAGCCACGGTGAGG 4878
QY 5502 TCTCTGCCACTCTCTGGAGAGGACGAGGATGTAGCCCTGTAAACGGGACCGGGTGTC 5561
DB 4879 TCTCTGCCACTCTCTGGAGAGGACGAGGATGTAGCCCTGTAAACGGGACCGGGTGTC 4938
QY 5562 CAAGGACCAACCGGGGATGTGCTGCTGAGGAATCCCTCTGAGTATACCGTGGG 5621
DB 4939 CAAGGACCAACCGGGGATGTGCTGCTGAGGAATCCCTCTGAGTATACCGTGGG 4998
QY 5622 CAGAGGATGCACTGTTGACCGCTTGTGCTTGGGATCCCTCTGAGTATACCGTGGG 5681
DB 4999 CAGAGGATGCACTGTTGACCGCTTGTGCTTGGGATCCCTCTGAGTATACCGTGGG 5058
QY 5682 CTGCTGGGTGTGAATGAGCAGGGAAGACGCTCCACGCTTTCGATGTGTACGGGGACACA 5741
DB 5059 CTGCTGGGTGTGAATGAGCAGGGAAGACGCTCCACGCTTTCGATGTGTACGGGGACACA 5118
QY 5742 TTGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAACCCAGTGCT 5801
DB 5119 TTGGCCAGCAGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCGGGAACCCAGTGCT 5178
QY 5802 GCGCACCCTCAGCATGGGATACTGCCCCATCGATGCCATCTTTGAGCTGCTGACGGGC 5861
DB 5179 GCGCACCCTCAGCATGGGATACTGCCCCATCGATGCCATCTTTGAGCTGCTGACGGGC 5238
QY 5862 CGCGAGCACCCTGAGCTGTGCGCGCTTGGCGGTGTCCCGAGGCCACAGGTTGCCAG 5921
DB 5239 CGCGAGCACCCTGAGCTGTGCGCGCTTGGCGGTGTCCCGAGGCCACAGGTTGCCAG 5298
QY 5922 ACCGCTGGCTCAGGCTTGGCGGTCTGGGACTCTCTCATGTGTACGACAGCGGCTGCAGGC 5981
DB 5299 ACCGCTGGCTCAGGCTTGGCGGTCTGGGACTCTCTCATGTGTACGACAGCGGCTGCAGGC 5358
QY 5982 ACCTACAGCGGAGGGAACCAACGACGCTGGCAGCGCCCTGGCTGTTGGGACCCA 6041
DB 5359 ACCTACAGCGGAGGGAACCAACGACGCTGGCAGCGCCCTGGCTGTTGGGACCCA 5418
QY 6042 GCGGTGCTGTTCTGACGAGCGGACCAACAGGATGACCCACGCGCGCGCTTCCCTT 6101
DB 5419 GCGGTGCTGTTCTGACGAGCGGACCAACAGGATGACCCACGCGCGCGCTTCCCTT 5478
QY 6102 TGAACAGCCTTTTGGCCGTGGTGGGAGGCGGCTTCAGTGTACCTCCCATPAGC 6161
DB 5479 TGAACAGCCTTTTGGCCGTGGTGGGAGGCGGCTTCAGTGTACCTCCCATPAGC 5538
QY 6162 ATGGAGAGTGTAGAGCTCTGCTCGCGCTAGCCATCATGTTGATGGCGGTTCCGC 6221
DB 5539 ATGGAGAGTGTAGAGCTCTGCTCGCGCTAGCCATCATGTTGATGGCGGTTCCGC 5598
QY 6222 TGCTTGGCAGCCCGCAACATCTCAAGGCGAGATTTCGCGGGGTACACACTGACCCCTG 6281
DB 5599 TGCTTGGCAGCCCGCAACATCTCAAGGCGAGATTTCGCGGGGTACACACTGACCCCTG 5658
QY 6282 CGGCTGCCCGCGCAAGGTCCCAAGCGGACGCGGCTTCTGTGGCGCGGAGTTCCCTGGG 6341
DB 5659 CGGCTGCCCGCGCAAGGTCCCAAGCGGACGCGGCTTCTGTGGCGCGGAGTTCCCTGGG 5718

QY 6342 TCGAGCTGCGGAGGCACATGAGGCCCGCTCGCTTCCAGCTGCCCGGAGGGCGC 6401
DB 5719 TCGAGCTGCGGAGGCACATGAGGCCCGCTCGCTTCCAGCTGCCCGGAGGGCGC 5778
QY 6402 TGGCCCTTGGCGCGCTTTTGGAGAGCTGGCGGTGCAAGGCGCAGACACGCGGTGGAG 6461
DB 5779 TGGCCCTTGGCGCGCTTTTGGAGAGCTGGCGGTGCAAGGCGCAGACACGCGGTGGAG 5838
QY 6462 GACTTTTTCGTGAGCCAGACGATGCTGGAGGAGTATCTTGTACTTCTCCAAAGGACCAG 6521
DB 5839 GACTTTTTCGTGAGCCAGACGATGCTGGAGGAGTATCTTGTACTTCTCCAAAGGACCAG 5898
QY 6522 GGAAGGACGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCCA 6581
DB 5899 GGAAGGACGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCCA 5958
QY 6582 GGCCTCGAGCACCCCAACCGCTGACCCAGTTCCTCGATGACCTAGCACTGCCGAGACT 6641
DB 5959 GGCCTCGAGCACCCCAACCGCTGACCCAGTTCCTCGATGACCTAGCACTGCCGAGACT 6018
QY 6642 GTCCTCTGAGCCTCCCTCCCTCGCGGGCGCGGGGAGGCCCTGGGAATGGCAAGGCCAA 6701
DB 6019 GTCCTCTGAGCCTCCCTCCCTCGCGGGCGCGGGGAGGCCCTGGGAATGGCAAGGCCAA 6078
QY 6702 GGTAGAGTGTAGGAGCCTGACTCAGGCTGGCAGAGGGGCTGGTGGCCTGGAGAAA 6761
DB 6079 GGTAGAGTGTAGGAGCCTGACTCAGGCTGGCAGAGGGGCTGGTGGCCTGGAGAAA 6138
QY 6762 TAAAGAGAGGCTGGAGAGAGCCGCTGCTGGTGGAAA 6797
DB 6139 TAAAGAGAGGCTGGAGAGAGCCGCTGGTGGTGGAAA 6174

RESULT 13

US-10-114-270-175
; Sequence 175, Application US/10114270
; Publication No. US20040030110A1

GENERAL INFORMATION:

; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerkhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Beha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086


```

; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 175
; LENGTH: 6327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6178)
; US-10-114-270-175

```

Query Match	88.7%;	Score 6036.8;	DB 16;	Length 6327;
Best Local Similarity	95.9%;	Prod. No. 0;		
Matches 6315;	Conservative 12;	Mismatches 261;	Gaps 1;	
Qy	210	ATGSCCTTCGACACACGCTCATGCTGCTCTCGAAGAAATTCATGTAATCGCGGAGA	269	
Db	1	ATGSCCTTCGACACACGCTCATGCTGCTCTCGAAGAAATTCATGTAATCGCGGAGA	60	
Qy	270	CAGCCGGTCCAGCTCCTGGTCCGAATGCTGTGGCCTCTCTTCTCTTCTTCATCTGGTG	329	
Db	61	CAGCCGGTCCAGCTCCTGGTCCGAATGCTGTGGCCTCTCTTCTCTTCTTCATCTGGTG	120	
Qy	330	GCTGTTCCGCACATCCCAACCCGCCCTCGAGACACCATGATGACATGCGCACTCCCGAACAAGCCA	389	
Db	121	GCTGTTCCGCACATCCCAACCCGCCCTCGAGACACCATGATGATGCCACTTCCCGAACAAGCCA	180	
Qy	390	CTGCCATCGCGGGACACGTCGCCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAACACC	449	
Db	181	CTGCCATCGCGGGACACGTCGCCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAACACC	240	
Qy	450	TGCTTTCCGACGTGACACCGGGCGAGAGCCCGGGCGCTTGAGCAACTTCAGACGTCC	509	
Db	241	TGCTTTCCGACGTGACACCGGGCGAGAGCCCGGGCGCTTGAGCAACTTCAGACGTCC	300	
Qy	510	CTGCTCTCCCGGCTGTAGCCGATGCCGCACTGTGCTGGAGAGGGCCAGTGCCCAAGG	569	
Db	301	CTGCTCTCCCGGCTGTAGCCGATGCCGCACTGTGCTGGAGAGGGCCAGTGCCCAAGG	360	
Qy	570	ACGCTGCTGGCCTTAGGGAAGCTGATCGCCAACGCTGAGGGGTGCACGACGACGCGCCAG	629	
Db	361	ACGCTGCTGGCCTTAGGGAAGCTGATCGCCAACGCTGAGGGGTGCACGACGACGCGCCAG	420	
Qy	630	CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	689	
Db	421	CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCATGCTGGATGTGCGGAGCTGCTG	480	
Qy	690	ACGTCACTGTGTCGCAACGGAATCCTCGGGGTGGCACTGGGCCAAGCCACGAGGCCCTTG	749	
Db	481	ACGTCACTGTGTCGCAACGGAATCCTCGGGGTGGCACTGGGCCAAGCCACGAGGCCCTTG	540	
Qy	750	CACAGCTGTGTGAGGCCGCTGAGGACCTGGCCCCAGGAGCTCTCGGCGCTGCGAGCCTG	809	
Db	541	CACAGCTGTGTGAGGCCGCTGAGGACCTGGCCCCAGGAGCTCTCGGCGCTGCGAGCCTG	600	

QY	810	GTGAGCTTTCGGGCACTGCTGCAGAGACCCCGAGGGACAGCGGCCCCCTTGGAGTTGCTG	869
DB	601	GTGAGCTTTCGGGCACTGCTGCAGAGACCCCGAGGGACAGCGGCCCCCTTGGAGTTGCTG	660
QY	870	TCAGAGCCCTCTGCACTGTTCAGGTGCAGGGACCTAGCAGCACACGTGGGCCCCCTCCCTCAACTGG	929
DB	661	TCAGAGCCCTCTGCACTGTTCAGGTGCAGGGACCTAGCAGCACACGTGGGCCCCCTCCCTCAACTGG	720
QY	930	TACGAGCTTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGAAATCCGCCCTGCCAGAC	989
DB	721	TACGAGCTTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGAAATCCGCCCTGCCAGAC	780
QY	990	AGCAGCTGAGCCCCCGCCTGCTGGAGCTGATTTGGAGCCCTTGGACAGCACACCCGCTGTGTC	1049
DB	781	AGCAGCTGAGCCCCCGCCTGCTGGAGCTGATTTGGAGCCCTTGGACAGCACACCCGCTGTGTC	840
QY	1050	CGCCTGCTTCGGAGACGCTTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGACACAGAT	1109
DB	841	CGCCTGCTTCGGAGACGCTTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGACACAGAT	900
QY	1110	ACACCTTTTACCCGGAGCTCATGGGCCAGGTGAAACCGGACCTTCGAGGAGCTCACCCCTG	1169
DB	901	ACACCTTTTACCCGGAGCTCATGGGCCAGGTGAAACCGGACCTTCGAGGAGCTCACCCCTG	960
QY	1170	CTGAGGATGTCCGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAAC	1229
DB	961	CTGAGGATGTCCGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAAC	1020
QY	1230	GACAGTTCCAAATGTGGCCATGCTGCAGCGGCTCTTCGAGATGCAGATGAGGAAGAGG	1289
DB	1021	GACAGTTCCAAATGTGGCCATGCTGCAGCGGCTCTTCGAGATGCAGATGAGGAAGAGG	1080
QY	1290	CAGCCACAGACTGGAGSCGGGACCAATGGAGGCCCTCGATCTTCTTGGACCCCTGGG	1349
DB	1081	CAGCCACAGACTGGAGSCGGGACCAATGGAGGCCCTCGATCTTCTTGGACCCCTGGG	1140
QY	1350	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCACTGTGGGCAAGCTG	1409
DB	1141	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCACTGTGGGCAAGCTG	1200
QY	1410	GGCCGAGTAGCAGAGTGCCTGTCTTTGGACAGACTGGAGGCGGCACTCTCAGAGGAGCC	1469
DB	1201	GGCCGAGTAGCAGAGTGCCTGTCTTTGGACAGACTGGAGGCGGCACTCTCAGAGGAGCC	1360
QY	1470	CTGTGTCGGGGCCCTGCACCTGCTCGGGACATCGATTTCTGGGCGGCGCTGCTTC	1529
DB	1261	CTGTGTCGGGGCCCTGCACCTGCTCGGGACATCGATTTCTGGGCGGCGCTGCTTC	1320
QY	1530	TTGGGACCTGAGGACTCTTCAGACCCCAAGAGCACTTGGGCACTTGGGCGGCGGCGG	1589
DB	1321	TTGGGACCTGAGGACTCTTCAGACCCCAAGAGCACTTGGGCACTTGGGCGGCGGCGG	1380
QY	1590	CAGTGGCATCAAAATTCGCATGGAATTTAGCGTGGTCAACGAGGACCAATAGATCAGG	1649
DB	1381	CAGTGGCATCAAAATTCGCATGGAATTTAGCGTGGTCAACGAGGACCAATAGATCAGG	1440
QY	1650	GACAGTTTGGGACCTTGGCCAGCGGCGGACCCCTTGCACGACCTTGGCTACGTGTGG	1709
DB	1441	GACAGTTTGGGACCTTGGCCAGCGGCGGACCCCTTGCACGACCTTGGCTACGTGTGG	1500
QY	1710	GGCGCTTCTGTGTACCTGCAAGACCTGTGGAGCGTGCAGCGCTCGCGTGTCTACGCGG	1769
DB	1501	GGCGCTTCTGTGTACCTGCAAGACCTGTGGAGCGTGCAGCGCTCGCGTGTCTACGCGG	1560
QY	1770	GCCAAACCCCGGCGGCTTACTCTGCAGAGATGCCCTATCCGCTGCTATGTGGACGAC	1829
DB	1561	GCCAAACCCCGGCGGCTTACTCTGCAGAGATGCCCTATCCGCTGCTATGTGGACGAC	1620
QY	1830	GTGTTCTGCTGTGTGAGCGGTCGCTGCGCTCTTCTTGCAGCTGGCGCTGGATCTTAC	1889
DB	1621	GTGTTCTGCTGTGTGAGCGGTCGCTGCGCTCTTCTTGCAGCTGGCGCTGGATCTTAC	1680

Qy	1890	TCCGTGACACTGACAGTGAAGCGCGTGTGCGGGAGAAAGAGACGCGGCTGCGGGACACC	1949
Db	1681	TCCGTGACACTGACAGTGAAGCGCGTGTGCGGGAGAAAGAGACGCGGCTGCGGGACACC	1740
Qy	1950	ATCGCGCCATGGGCTCAGCGCGGCTGCTCTGGCTAGGCTGGTTCCTCAGCTGCCTC	2009
Db	1741	ATCGCGCCATGGGCTCAGCGCGGCTGCTCTGGCTAGGCTGGTTCCTCAGCTGCCTC	1800
Qy	2010	GGGCGCTTCCTGCTCAGCGCGGCTGCTCTGGTTCCTGGTTCCTCAGCTGGGGACATCCCTC	2069
Db	1801	GGGCGCTTCCTGCTCAGCGCGGCTGCTCTGGTTCCTGGTTCCTCAGCTGGGGACATCCCTC	1860
Qy	2070	CCCTACAGCCACCGCGCGTGGTTCCTCTCTTTGGCAGCTTCGCGGTGGCCACGGTG	2129
Db	1861	CCCTACAGCCACCGCGCGTGGTTCCTCTCTTTGGCAGCTTCGCGGTGGCCACGGTG	1920
Qy	2130	ACCCAGAGCTTCCTGCTCAGGGCTTCCTCTCCGCGCCAACTGGCTGCGGCTGCGGC	2189
Db	1921	ACCCAGAGCTTCCTGCTCAGGGCTTCCTCTCCGCGCCAACTGGCTGCGGCTGCGGC	1980
Qy	2190	GGCTGCGCTACTTCCTCCCTACTACCTGCCCTACGTGCTGTGTGGCTTCGGCGGACCGG	2249
Db	1981	GGCTGCGCTACTTCCTCCCTACTACCTGCCCTACTGTCTGTGTGGCTTCGGCGGACCGG	2040
Qy	2250	CTGCGCGCGGCTGGCGCGCTGGCGGAGCTGCTGTGCGCCGTGGGCTTCGGCTTCGGC	2309
Db	2041	CTGCGCGCGGCTGGCGCGCTGGCGGAGCTGCTGTGCGCCGTGGGCTTCGGCTTCGGC	2100
Qy	2310	TGGAGAGCTTGCTCTGTGTGGAGGACAGGGGAGGGCGCGCAGTGGCACACGTGGGC	2369
Db	2101	TGGAGAGCTTGCTCTGTGTGGAGGACAGGGGAGGGCGCGCAGTGGCACAACTGGGC	2160
Qy	2370	ACCGGCGCTACGGCAGACGCTCTTCAGCCTTGCGCCAGGCTCTCGGCTTCCTGCTGTGGAC	2429
Db	2161	ACCGGCGCTACGGCAGACGCTCTTCAGCCTTGCGCCAGGCTCTCGGCTTCCTGCTGTGGAC	2220
Qy	2430	GCGGCGCTACTAGGCGCTCGCACCTGTGTACTGTGAAGCTGTGTGCCAGGCCAGTACGGG	2489
Db	2221	GCGGCGCTACTAGGCGCTCGCACCTGTGTACTGTGAAGCTGTGTGCCAGGCCAGTACGGG	2280
Qy	2490	ATCCCTGAAACGATGGAAATTTTCTTTTCGGAGGAGCTACTGTGTGCGGACCTCGGCC	2549
Db	2281	ATCCCTGAAACGATGGAAATTTTCTTTTCGGAGGAGCTACTGTGTGCGGACCTCGGCC	2340
Qy	2550	AAGAGTCCAGCCCTTCGCCCCACCCGCTGAGACCCAAAGGTGCTGTGTGAAGAGGACACCG	2609
Db	2341	AAGAGTCCAGCCCTTCGCCCCACCCGCTGAGACCCAAAGGTGCTGTGTGAAGAGGACACCG	2400
Qy	2610	CCGCGCTGAGTCTCTGGCGTATCCGTTTCGACGCTGGAGAGCGCTTCCTCTGGAAGCCCG	2669
Db	2401	CCGCGCTGAGTCTCTGGCGTATCCGTTTCGACGCTGGAGAGCGCTTCCTCTGGAAGCCCG	2460
Qy	2670	CAGCCAGCTTCGCGGGGCTCAGCCTGGACTTCTAACAGGGCCACATCAACGCTTCCTG	2729
Db	2461	CAGCCAGCTTCGCGGGGCTCAGCCTGGACTTCTAACAGGGGCCACATCAACGCTTCCTG	2520
Qy	2730	GGCCACAACGGGCGCGAGACACACCCCTGTCCATCTTGTAGTGGGCTCTTCCACCC	2789
Db	2521	GGCCACAACGGGCGCGAGACACACCCCTGTCCATCTTGTAGTGGGCTCTTCCACCC	2580
Qy	2790	AGTGGTGGCTCTGCTTTCATCTCTGGGCGACGACGCTCGGCTCCAGCATGGCGGCTCCGG	2849
Db	2581	AGTGGTGGCTCTGCTTTCATCTCTGGGCGACGACGCTCGGCTCCAGCATGGCGGCTCCGG	2640
Qy	2850	CCCCACTGTGGCGCTGTGTCTCAGTAAACAGTGTCTTTTGACATGTCTGAACGTGGACGAG	2909
Db	2641	CCCCACTGTGGCGCTGTGTCTCAGTAAACAGTGTCTTTTGACATGTCTGAACGTGGACGAG	2700
Qy	2910	CAGCTGTGTCTATATGGCGGCTGAAGGCTCTGAGTCCGCTGTGTGTGGGCCCGGACGAG	2969
Db	2701	CAGCTGTGTGTCTATATGGCGGCTGAAGGCTCTGAGTCCGCTGTGTGTGGGCCCGGACGAG	2760
Qy	2970	GACCGTCTGTGTGACAGGATGTGGGCTGGTCTTCCAAAGCAGAGTGTGCAGACTCGCCACCTC	3029

2761		GACCGTCTGCTGCAGGATGTGGGGCTGGTCTCAAGACAGAGTGTGCAGACTTCGCCCACTC	2820
3030		TCCTGTGGGATGCAACGGAAGCTGTCCGTGGGCATTTGCTTTGTGGGCGGTCTCCAAAGTT	3089
2821		TCCTGTGGGATGCAACGGAAGCTGTCCGTGGGCATTTGCTTTGTGGGCGGTCTCCAAAGTT	2880
3090		GTTATCTCGACGAGCTTACGGCTGGCTGGATCTCTGCTTCCGCCGCGTATTTGGGAG	3149
2881		GTTATCTCGACGAGCTTACGGCTGGCTGGATCTCTGCTTCCGCCGCGTATTTGGGAG	2940
3150		CTGCTGCTCAAAATACCGAAGGTTCGACGCTGATCTCTCCACCCACCACTTGGATGAG	3209
2941		CTGCTGCTCAAAATACCGAAGGTTCGACGCTGATCTCTCCACCCACCACTTGGATGAG	3000
3210		CGAGAGCTCTGGGAGACCGTGTGGCTGTGTGGCAGGTGGCGCTTGTGCTGTCTGGC	3269
3001		CGAGAGCTCTGGGAGACCGTGTGGCTGTGTGGCAGGTGGCGCTTGTGCTGTCTGGC	3060
3270		TCCCACTCTTTCTGCGCCGTCACTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC	3329
3061		TCCCACTCTTTCTGCGCCGTCACTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC	3120
3330		CGCTGCCCCGTGACCAACAATGAGAAGCTGACATGACATGGAGGGCAGTGTGGACAC	3389
3121		CGCTGCCCCGTGACCAACAATGAGAAGCTGACATGACATGGAGGGCAGTGTGGACAC	3180
3390		AGCGAGGAAAAGAGAATGCGACGAGGCGACGAGGTGGGCACCTTCCTACGCTCTGGCC	3449
3181		AGCGAGGAAAAGAGAATGCGACGAGGCGACGAGGTGGGCACCTTCCTACGCTCTGGCC	3240
3450		CTGTGTACAGCACTGGGTGCCCCGGGGCACGGCTGGTGGAGGAGCTGSCACACAGAGCTGGTG	3509
3241		CTGTGTACAGCACTGGGTGCCCCGGGGCACGGCTGGTGGAGGAGCTGSCACACAGAGCTGGTG	3300
3510		CTGTGTGTCCTTACACGGGTGCCATGACGGCAGGTTTGGCCACATCTTCCGAGAGCTTA	3569
3301		CTGTGTGTCCTTACACGGGTGCCATGACGGCAGGTTTGGCCACATCTTCCGAGAGCTTA	3360
3570		GACACGGCTTGGCGGAGCTGAGGCTCACTTGGCTACGGGATCTCCGACACACAGAGCTCGAG	3629
3361		GACACGGCTTGGCGGAGCTGAGGCTCACTTGGCTACGGGATCTCCGACACACAGAGCTCGAG	3420
3630		GAGATCTTCTGAAGGTGGTGGAGGAGTGTGCTCGGACACAGATATGAGGAGTGGCAGC	3689
3421		GAGATCTTCTGAAGGTGGTGGAGGAGTGTGCTCGGACACAGATATGAGGAGTGGCAGC	3480
3690		TGGGGGAGACCTTATGACAGGCATTTGCTGGCTAGAGCTTAACCTCGGCTCAAGATG	3749
3481		TGGGGGAGACCTTATGACAGGCATTTGCTGGCTAGAGCTTAACCTCGGCTCAAGATG	3540
3750		CCGCCACAGGAGACAGCGCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCCAGACACTGAC	3809
3541		CCGCCACAGGAGACAGCGCTGGAGAACGGGGAACAGCTGGGTGAGCCCGCCAGACACTGAC	3600
3810		CAGGGCTCTGGGCCAGACGCGTGGGCGGGTACAGGGCTGGGGCACTGACCCGCCAGCAG	3869
3601		CAGGGCTCTGGGCCAGACGCGTGGGCGGGTACAGGGCTGGGGCACTGACCCGCCAGCAG	3660
3870		CTTCAGGCCCTGCTTCTCAAGCGCTTTCGTGTGGCCCGCAGCGCCGGCGGCTGTTC	3929
3661		CTTCAGGCCCTGCTTCTCAAGCGCTTTCGTGTGGCCCGCAGCGCCGGCGGCTGTTC	3720
3930		GGCCAGATGCTGCTCTGCTGCTTGTGGGCTGGCCCTCGTGTTCAGGCTCATCTGTG	3989
3721		GGCCAGATGCTGCTCTGCTGCTTGTGGGCTGGCCCTCGTGTTCAGGCTCATCTGTG	3780
3990		CTCTCTTTGGGCATACCCGGCTCTCGCGCTCAGTCCCAACCATGACGCTGCTCAGGTG	4049
3781		CTCTCTTTGGGCATACCCGGCTCTCGCGCTCAGTCCCAACCATGACGCTGCTCAGGTG	3840
4050		TCTCTTCTTCAGTGAGGACGCCCAAGGGGACCTTGGAGCTGCCCGGCTGCTCGAGGCGCTG	4109

Db	3499	GGACTGGAGAGCCCCAGTGCAGCATAGTCCACAGGTTCTCGGCACACAGAAAGTTTCCT	3558
Qy	4182	GCTGAAGTGGCCAAAGGCTTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCC	4241
Db	3559	GCTGAAGTGGCCAAAGGCTTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCC	3618
Qy	4242	TGCCAGTGTAGCAGCCCGGTGGCCCGCGCTGCTGCTGCCGACTGCCCCGGCTGCAGCTGGT	4301
Db	3619	TGCCAGTGTAGCAGCCCGGTGGCCCGCGCTGCTGCTGCCGACTGCCCCGGCTGCAGCTGGT	3678
Qy	4302	GGTCCCCCTCCGCCCCCAGGCAAGTGAACCGGCTCTGGGAAAGTGGTTCAAGAACCTTGACAGGC	4361
Db	3679	GGTCCCCCTCCGCCCCCAGGCAAGTGAACCGGCTCTGGGAAAGTGGTTCAAGAACCTTGACAGGC	3738
Qy	4362	CGGAACCTGTCTGACTTCTCTGTTCAAGACCTACCCGCGCTGGTGGCCAGAGGCTTGAAG	4421
Db	3739	CGGAACCTGTCTGACTTCTCTGTTCAAGACCTACCCGCGCTGGTGGCCAGAGGCTTGAAG	3798
Qy	4422	ACTAAGAAGTGGGTGAATGAGTCAAGTACGAGAGCTTCTCGCTGGGGGCCGAGACCCA	4481
Db	3799	ACTAAGAAGTGGGTGAATGAGTCAAGTACGAGAGCTTCTCGCTGGGGGCCGAGACCCA	3858
Qy	4482	GGCCTGCCCCCTCGGCCCAAGAGTTGGCCCGCTCAAGTGGAGAGTTGTGGCGCTGCTGAGT	4541
Db	3859	GGCCTGCCCCCTCGGCCCAAGAGTTGGCCCGCTCAAGTGGAGAGTTGTGGCGCTGCTGAGT	3918
Qy	4542	CCCTGCTGGCGGGGCCCTCGACGCTGTCTGTAAGAACCTCACAGCTGGGCTCACAGC	4601
Db	3919	CCCTGCTGGCGGGGCCCTCGACGCTGTCTGTAAGAACCTCACAGCTGGGCTCACAGC	3978
Qy	4602	CTGATGCTCAGGACAGTCTCAAGATCTGGTTCAAGAACAGGCTGGCACTCCATGGTG	4661
Db	3979	CTGATGCTCAGGACAGTCTCAAGATCTGGTTCAAGAACAGGCTGGCACTCCATGGTG	4038
Qy	4662	GCCTTTGTCAACCGAGCGCAACAGCAATCCTCGTGTCTCACTGCCCCCGAGCCCGGCC	4721
Db	4039	GCCTTTGTCAACCGAGCGCAACAGCAATCCTCGTGTCTCACTGCCCCCGAGCCCGGCC	4098
Qy	4722	CGGCAGCCCCACAGATFACACACACTCAACACCCCTTGAACTCAACAGAGAGCAGCTG	4781
Db	4099	CGGCAGCCCCACAGATFACACACACTCAACACCCCTTGAACTCAACAGAGAGCAGCTG	4158
Qy	4782	TCTGAGCTGACATGAGGCTCTCGGTGAGGCTCTCGTCTCCATCTGCTGTGCTTTT	4841
Db	4159	TTTGAGGCTGACATGAGGCTCTCGGTGAGGCTCTCGTCTCCATCTGCTGTGCTTTT	4218
Qy	4842	GCCATGTCTTTGTCCCGGCCAGCTTCACTCTTTGTCTCTCAATTGAGGAGCGAGTCAACCGA	4901
Db	4219	GCCATGTCTTTGTCCCGGCCAGCTTCACTCTTTGTCTCTCAATTGAGGAGCGAGTCAACCGA	4278
Qy	4902	GCCAGACCTGAGCTCATGGGGGGCTGTCCCGACCCCTCTACTGGCTTGGCAACTTT	4961
Db	4279	GCCAGACCTGAGCTCATGGGGGGCTGTCCCGACCCCTCTACTGGCTTGGCAACTTT	4338
Qy	4962	CTCTGGACATGTAACTTTTGGTGGCCAGATGATCGTGTGCTCATCTTTCTGGCC	5021
Db	4339	CTCTGGACATGTAACTTTTGGTGGCCAGATGATCGTGTGCTCATCTTTCTGGCC	4398
Qy	5022	TTCCAGCAGAGGCGATATGTGGCCCTCGCCAACTGCTGCTCTCTGCTGTGTGCTACTA	5081
Db	4399	TTCCAGCAGAGGCGATATGTGGCCCTCGCCAACTGCTGCTCTCTGCTGTGTGCTACTA	4458
Qy	5082	CTGTATGGCTGTGCTGATFACACCGCTCATGTACCCAGCTCTCTTCTCTCCGTGCC	5141
Db	4459	CTGTATGGCTGTGCTGATFACACCGCTCATGTACCCAGCTCTCTTCTCTCCGTGCC	4518
Qy	5142	AGCACAGCTATGTGGTCTCACTGCATAAACCTCTTTATTGGCATCAATGAAGCATG	5201
Db	4519	AGCACAGCTATGTGGTCTCACTGCATAAACCTCTTTATTGGCATCAATGAAGCATG	4578
Qy	5202	GCCACCTTTGTGCTTGAAGCTTCTCTGATCAGAAGCTGCAGGAGGTGAGCCGATCTTG	5261

Db	4579	GCCACCTTTTGTGCTTGAAGCTTCTCTGATCAGAAGCTGCAGGAGGTGAGCCGATCTTG	4638
Qy	5262	AAACAGGCTTTCCTTATCTTCCCCCACTTCTGCTTGGCGGGGCTCATTTGACATGGTG	5321
Db	4639	AAACAGGCTTTCCTTATCTTCCCCCACTTCTGCTTGGCGGGGCTTATTTGACATGGTG	4698
Qy	5322	CGGAACACAGCCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAAGTTCACAGTCAACC	5381
Db	4699	CGGAACACAGCCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAAGTTCACAGTCAACC	4758
Qy	5382	CTGCGTGGGAGGTGCTGCGCAAGAACCTTCTTGGCCATGCTGATACAGGGGCCCTCTTC	5441
Db	4759	CTGCGTGGGAGGTGCTGCGCAAGAACCTTCTTGGCCATGCTGATACAGGGGCCCTCTTC	4818
Qy	5442	CTTCTCTTCACTACTGCTGAGCAGCCGAAAGCAACTCTCTGCCACAGCCACGGGTGAGG	5501
Db	4819	CTTCTCTTCACTACTGCTGAGCAGCCGAAAGCAACTCTCTGCCACAGCCACGGGTGAGG	4878
Qy	5502	TCTCTGCCACTCTCTGGGAGAGGACGAGGATGTAGCCCGTGAACGGGAGCGGTGGTC	5561
Db	4879	TCTCTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAACGGGAGCGGTGGTC	4938
Qy	5562	CAAGGAGCCACCCAGGGGGATGTGGTCTGAGGAACTTGAACAAAGTATACCTGGG	5621
Db	4939	CAAGGAGCCACCCAGGGGGATGTGGTCTGAGGAACTTGAACAAAGTATACCTGGG	4998
Qy	5622	CAGAGATGCCAGCTTTCGACCGCTTGTGCTGGGGATTTCCCTCTGGTGTGTTTTGGG	5681
Db	4999	CAGAGATGCCAGCTTTCGACCGCTTGTGCTGGGGATTTCCCTCTGGTGTGTTTTGGG	5058
Qy	5682	CTGCTGGGTGTGAATCGAGCAGGGAAGACGTCCACGTTTCGCATGTGTGACGGGGACACA	5741
Db	5059	CTGCTGGGTGTGAATCGAGCAGGGAAGACGTCCACGTTTCGCATGTGTGACGGGGACACA	5118
Qy	5742	TTGGCCAGCAGGGCGAGGCTGTGTGGCAGGCAACAGCGTGGCCCGGGAACCCAGTGCT	5801
Db	5119	TTGGCCAGCAGGGCGAGGCTGTGTGGCAGGCAACAGCGTGGCCCGGGAACCCAGTGCT	5178
Qy	5802	GCGCACTCAGCATGGGATGCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGC	5861
Db	5179	GCGCACTCAGCATGGGATGCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGC	5238
Qy	5862	CGCGAGCACTGGAGCTGTGTGGCGCCCTGCGCGGTGTCCCGAGGCCCAGAGTTGCCAG	5921
Db	5239	CGCGAGCACTGGAGCTGTGTGGCGCCCTGCGCGGTGTCCCGAGGCCCAGAGTTGCCAG	5298
Qy	5922	ACCGTGGCTCAGGCTGTGGGACTCTCATGTGTACGACAGCGGCTTCGAGGC	5981
Db	5299	ACCGTGGCTCAGGCTGTGGGACTCTCATGTGTACGACAGCGGCTTCGAGGC	5358
Qy	5982	ACCTACAGCGGAGGGAACAAACGCAAGCTTGGCGACGGCCCTGGCGTGGTGGGGACCCA	6041
Db	5359	ACCTACAGCGGAGGGAACAAACGCAAGCTTGGCGACGGCCCTGGCGTGGTGGGGACCCA	5418
Qy	6042	GCGTGGTGTCTGACAGCGCGACACAGGATGGAACCCAGCGCGCGGCTTCCTT	6101
Db	5419	GCGTGGTGTCTGACAGCGCGACACAGGATGGAACCCAGCGCGCGGCTTCCTT	5478
Qy	6102	TGGAACAGGCTTTTGGCCGCTGGTGGCGGAGGGCGTTCAGTGTATGCTCACCTCCATAGC	6161
Db	5479	TGGAACAGGCTTTTGGCCGCTGGTGGCGGAGGGCGTTCAGTGTATGCTCACCTCCATAGC	5538
Qy	6162	ATGGAGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGGTGAATGGGCGGTTCCGC	6221
Db	5539	ATGGAGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGGTGAATGGGCGGTTCCGC	5598
Qy	6222	TGCTGGGAGCGCCGCAACATCTCAAGGCGAGATTCGCGGGGTCAACACTGACCCCTG	6281
Db	5599	TGCTGGGAGCGCCGCAACATCTCAAGGCGAGATTCGCGGGGTCAACACTGACCCCTG	5658
Qy	6282	CGGGTCCCGCGCAAGGTTCCAGCGCGGAGCGGCTTCTGTTGCGCGCCGAGTTCCCTGGG	6341
Db	5659	CGGGTCCCGCGCAAGGTTCCAGCGCGGAGCGGCTTCTGTTGCGCGCCGAGTTCCCTGGG	5718

1590 QY CAGTGGCATCAAAATCCGATGACATTGACGTGGTACAGGACCAATAAGATCAGG 1649
1381 Db CAGTGGCATCAAAATCCGATGACATTGACGTGGTACAGGACCAATAAGATCAGG 1440
1650 QY GACAGGTTTTGGGACCCCTGGCCACAGCCGAGACCCCTTGACCGACCTGCGCTACGTGTGG 1709
1441 Db GACAGGTTTTGGGACCCCTGGCCACAGCCGAGACCCCTTGACCGACCTGCGCTACGTGTGG 1500
1710 QY GCGGCTTTGTTACCTGCAAGACCTGGTGGAGCGTGCAGCGCTCCGCTGCTCAGCGGC 1769
1501 Db GCGGCTTTGTTACCTGCAAGACCTGGTGGAGCGTGCAGCGCTCCGCTGCTCAGCGGC 1560
1770 QY GCCAACCCCGGGCCGCTCTACTCTGACAGATGCCCTATCCGTGCTATGTGGAGCAG 1829
1561 Db GCCAACCCCGGGCCGCTCTACTCTGACAGATGCCCTATCCGTGCTATGTGGAGCAG 1620
1830 QY GTGTTCTGCTGTGCTGAGCCGCTGCTGCTGCTCTTCTGACGCTGGCTGGATCTAC 1889
1621 Db GTGTTCTGCTGTGCTGAGCCGCTGCTGCTGCTCTTCTGACGCTGGCTGGATCTAC 1680
1890 QY TCCGTGACATGACAGTGAAGCCGCTGGTGGGAGAAAGAGACGCGGCTGCGGGACACC 1949
1681 Db TCCGTGACATGACAGTGAAGCCGCTGGTGGGAGAAAGAGACGCGGCTGCGGGACACC 1740
1950 QY ATGCGCGCATGGGCTCAGCGCGGCTGCTCTGCTAGGCTGGTTCCTCAGCTGCGCTC 2009
1741 Db ATGCGCGCGTGGGCTCAGCGCGGCTGCTCTGCTAGGCTGGTTCCTCAGCTGCGCTC 1800
2010 QY GGGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTCAAGCTGGGGGACATCCTC 2069
1801 Db GGGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTCAAGCTGGGGGACATCCTC 1860
2070 QY CCCTAGACCAACCGCGGCTGGTCTTCTGCTTTGAGCAGCTTGGCGGTGGCCAGCGTG 2129
1861 Db CCCTAGACCAACCGCGGCTGGTCTTCTGCTTTGAGCAGCTTGGCGGTGGCCAGCGTG 1920
2130 QY ACCAGAGCTTCCTGCTCAGCGCTTCTTCTCCGCGCCAACTGGCTGGCGGCTCGGC 2189
1921 Db ACCAGAGCTTCCTGCTCAGCGCTTCTTCTCCGCGCCAACTGGCTGGCGGCTCGGC 1980
2190 QY GGCCTGCGCTTCTTCTCCTCTACTGCTGCTGCTGCTGCTGCTGGCGGACCGG 2249
1981 Db GGCCTGCGCTTCTTCTCCTCTACTGCTGCTGCTGCTGCTGCTGGCGGACCGG 2040
2250 QY CTGCGCGGGTGGCGCGGTGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2309
2041 Db CTGCGCGGGTGGCGCGGTGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
2310 QY TGGAGAGCTTGGCTGCTGAGGAGCAGGGGAGGCGGCGGCGGAGCTGGGACAGCTGGGC 2369
2101 Db TGGAGAGCTTGGCTGCTGAGGAGCAGGGGAGGCGGCGGCGGAGCTGGGACAGCTGGGC 2160
2370 QY ACCCGGCTACGGCAGACGCTTCTCAGCTGGCGGCGAGGCTCTGGGCTTCTGCTGCTGGAC 2429
2161 Db ACCCGGCTACGGCAGACGCTTCTCAGCTGGCGGCGAGGCTCTGGGCTTCTGCTGCTGGAC 2220
2430 QY GCGGCGCTTACGCGCTTGGCACCTGTTA CTTGGAAGCTGTGTGCCAGGCGAGTACGGG 2489
2221 Db GCGGCGCTTACGCGCTTGGCACCTGTTA CTTGGAAGCTGTGTGCCAGGCGAGTACGGG 2280
2490 QY ATCCCTGAACCATGGAAATTTCTTTTTCGGAGGAGCTACTGTTGCGGACCTCGGCCCCC 2549
2281 Db ATCCCTGAACCATGGAAATTTCTTTTTCGGAGGAGCTACTGTTGCGGACCTCGGCCCCC 2340
2550 QY AAGAGTCCAGCCCTTGGCCACCCCGCTGGA CCCCAGGCTGTGTGTAAGAGGACACCG 2609
2341 Db AAGAGTCCAGCCCTTGGCCACCCCGCTGGA CCCCAGGCTGTGTGTAAGAGGACACCG 2400
2610 QY CCGGCGCTGAGTCTCGGCTATCCGTTGCGAGCTGCGAGGAGCGCTTCTTGAAGCCCG 2669
2401 Db CCGGCGCTGAGTCTCGGCTATCCGTTGCGAGCTGCGAGGAGCGCTTCTTGAAGCCCG 2460

2670 QY CAGCCAGCCCTGCGGGGCTCAGCTGGAGCTTCTACAGGGCCACATCAGCGCTTCTCTG 2729
2461 Db CAGCCAGCCCTGCGGGGCTCAGCTGGAGCTTCTACAGGGCCACATCAGCGCTTCTCTG 2520
2730 QY GGCCACAACGGGGCCGGCAAGACCAACCTGTGTCATCTTGAGTGGCTTCTTCCACCC 2789
2521 Db GGCCACAACGGGGCCGGCAAGACCAACCTGTGTCATCTTGAGTGGCTTCTTCCACCC 2580
2790 QY AGTGGTGGCTTGCCTTTCATCTTGGGGCCAGAGCTCGCTCCAGCATGGCGCCATCCGG 2849
2581 Db AGTGGTGGCTTGCCTTTCATCTTGGGGCCAGAGCTCGCTCCAGCATGGCGCCATCCGG 2640
2850 QY CCCCACTGGGGCTGCTGCTCTCAGTACAAACGCTGCTTTTGACATGCTGACCCGTGGACGAG 2909
2641 Db CCCCACTGGGGCTGCTGCTCTCAGTACAAACGCTGCTTTTGACATGCTGACCCGTGGACGAG 2700
2910 QY CACGTCTGTTTCTATGGGGCTGAAAGGCTTGAAGTCCGCTGTAGTGGGCCCCGAGCAG 2969
2701 Db CACGTCTGTTTCTATGGGGCTGAAAGGCTTGAAGTCCGCTGTAGTGGGCCCCGAGCAG 2760
2970 QY GACCGTCTGCTGACAGATGTGGGGCTGGTCTCCAAGCAGAGTGTGACAGCTCGGCACCTC 3029
2761 Db GACCGTCTGCTGACAGATGTGGGGCTGGTCTCCAAGCAGAGTGTGACAGCTCGGCACCTC 2820
3030 QY TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGGCTTTTGTGGCGGCTCCCAAGTT 3089
2821 Db TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGGCTTTTGTGGCGGCTCCCAAGTT 2880
3090 QY GTTATCTCTGACGAGCTTACGGCTGGCTGGATTCCTGTCTTCCCGCCCGCGTATTGGGAG 3149
2881 Db GTTATCTCTGACGAGCTTACGGCTGGCTGGATTCCTGTCTTCCCGCCCGCGTATTGGGAG 2940
3150 QY CTGCTCTCAATACACGAGAGTCCGACGCTGATCCTCTCCACACACACCTGGATGAG 3209
2941 Db CTGCTCTCAATACACGAGAGTCCGACGCTGATCCTCTCCACACACACCTGGATGAG 3000
3210 QY GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGCTGTGGC 3269
3001 Db GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGCTGTGGC 3060
3270 QY TCCCCACTTCTCTGCGCGCTACCTTGGGCTCCGGCTATCTACCTGACGCTGTGTGAAGGCC 3329
3061 Db TCCCCACTTCTCTGCGCGCTACCTTGGGCTCCGGCTATCTACCTGACGCTGTGTGAAGGCC 3120
3330 QY CGCTGCGCTTACACCAATGAGAGGCTGACACTGACATGAGGCGGAGTGGGACAC 3389
3121 Db CGCTGCGCTTACACCAATGAGAGGCTGACACTGACATGAGGCGGAGTGGGACAC 3180
3390 QY AGGACGAAAAGAAATGSCAGCCAGGCGCAGAGTCCGGACTCTCTCAGCTGTCTGGCC 3449
3181 Db AGGACGAAAAGAAATGSCAGCCAGGCGCAGAGTCCGGACTCTCTCAGCTGTCTGGCC 3240
3450 QY CTGGTACAGCACTTGGTGGCCCGGGCACA CCGCTGGTGGAGAGTGCACACGAGCTGGTG 3509
3241 Db CTGGTACAGCACTTGGTGGCCCGGGCACA CCGCTGGTGGAGAGTGCACACGAGCTGGTG 3300
3510 QY CTGGTCTGCCCTTACACGGGTGCCATGACGGGAGCTTCCGCCACTCTTCCGAGAGCTA 3569
3301 Db CTGGTCTGCCCTTACACGGGTGCCATGACGGGAGCTTCCGCCACTCTTCCGAGAGCTA 3360
3570 QY GACACGCGGTGGCGGAGCTGAGGCTCACTTGGCTAGCGGATCTCCGACACACGCTCGAG 3629
3361 Db GACACGCGGTGGCGGAGCTGAGGCTCACTTGGCTAGCGGATCTCCGACACACGCTCGAG 3420
3630 QY GAGATCTTCTGAAGTGGTGGAGGAGTGTGTGCGGACACAGATATGAGAGTATGGGAGC 3689
3421 Db GAGATCTTCTGAAGTGGTGGAGGAGTGTGTGCGGACACAGATATGAGAGTATGGGAGC 3480
3690 QY TGGGCGCAGCACCTATGACAGGCATTTGGCTAGAGTAACTCCCTGGGCTCAAGATG 3749
3481 Db TGGGCGCAGCACCTATGACAGGCATTTGGCTAGAGTAACTCCCTGGGCTCAAGATG 3540
3750 QY CCGCCACAGGAGACAGCGCTGGAGAACGGGAAACAGCTGGGTTCAGCCCCCAGAGACTGAC 3809

Db	3541		CCGCCACAGGAGACAGCGCTGGAGAACGGGGAAACCACTGGTTCAGCCCCAGAGACTGCAC	3600
Qy	3810		CAGGGCTCTGGGCCAGACGCGCTGGGCGGGGTACAGGGCTGGGCACCTGACCCCGCAGCAG	3869
Db	3601		CAGGGCTCTGGGCCAGACGCGCTGGGCGGGGTACAGGGCTGGGCACCTGACCCGCCAGCAG	3660
Qy	3870		CTCAGAGCCCTGTCTTCAAGCGCTTCTGCTTGGCCCGCCGACGCGCCGCGCTGTTC	3929
Db	3661		CTCAGAGCCCTGTCTTCAAGCGCTTCTGCTTGGCCCGCCGACGCGCGCGCTGTTC	3720
Qy	3930		GCCAGATCGTGTGCTGCTTGTGGGCTGCGCTTGTGGTTCAGGCTCATCGTG	3989
Db	3721		GCCAGATCGTGTGCTGCTTGTGGGCTGCGCTTGTGGTTCAGGCTCATCGTG	3780
Qy	3990		CCTCCTTTGGGCACCTACCCGGCTCTCGGGCTCAGTCCCAACCATGTACGGTGTCTCAGTG	4049
Db	3781		CCTCCTTTGGGCACCTACCCGGCTCTCGGGCTCAGTCCCAACCATGTACGGTGTCTCAGTG	3840
Qy	4050		TCCTTCTTCAGTGAGGACGCCACAGGGACCTCTGGACGTGCCGGCTGTCTGAGGCGCTG	4109
Db	3841		TCCTTCTTCAGTGAGGACGCCACAGGGACCTCTGGACGTGCCGGCTGTCTGAGGCGCTG	3900
Qy	4110		CTGCAGAGCAGGACTTGGAGGACGCCACAGGGACCTCTGGACGTGCCAGTATCTCCACAGGTTCTCGGCA	4169
Db	3901		CTGCAGAGCAGGACTTGGAGGACGCCACAGGGACCTCTGGACGTGCCAGTATCTCCACAGGTTCTCGGCA	3960
Qy	4170		CCAGAAGTCTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGACCTCCACAGAGTCT	4229
Db	3961		CCAGAAGTCTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGACCTCCACAGAGTCT	4020
Qy	4230		CCATCCCAAGCTGCCAGTGTAGCCAGCCGGTGTGCCGGCGCTGTGCTGCCGACTGCCCCG	4289
Db	4021		CCATCCCAAGCTGCCAGTGTAGCCAGCCGGTGTGCCGGCGCTGTGCTGCCGACTGCCCCG	4080
Qy	4290		GCTGCAGCTGGTGGTCCCTCCGCCACAGCAGTACCCGGCTCTGGGGAAGTGGTTCAG	4349
Db	4081		GCTGCAGCTGGTGGTCCCTCCGCCACAGCAGTACCCGGCTCTGGGGAAGTGGTTCAG	4140
Qy	4350		AACCTGACAGCCGGAACCTGTCTGACTTCTTGGTCAAGACCTACCCGGCGCTGTGGCGC	4409
Db	4141		AACCTGACAGCCGGAACCTGTCTGACTTCTTGGTCAAGACCTACCCGGCGCTGTGGCGC	4200
Qy	4410		CAGGGCTGAAGACTAAGAAAGTGGTGAATGAGGTACGAGGCTTCTCGCTGGGG	4469
Db	4201		CAGGGCTGAAGACTAAGAAAGTGGTGAATGAGGTACGAGGCTTCTCGCTGGGG	4260
Qy	4470		GGCGGAGACCCAGGCTGCCCTCGGCCAAGAGTTGGGCGCGCTCAGTGGAGGAGTTGTGG	4529
Db	4261		GGCGGAGACCCAGGCTGCCCTCGGCCAAGAGTTGGGCGCGCTCAGTGGAGGAGTTGTGG	4320
Qy	4530		GCCTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTCTGAAAACCTCACAGCC	4589
Db	4321		GCCTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTCTGAAAACCTCACAGCC	4380
Qy	4590		TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGCTGG	4649
Db	4381		TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGCTGG	4440
Qy	4650		CACCTCAATGGTGGCTTTGTTCACCCAGCCAGCAACCGCAATCTCCGTGTCTACCTGCC	4709
Db	4441		CACCTCAATGGTGGCTTTGTTCACCCAGCCAGCAACCGCAATCTCCGTGTCTACCTGCC	4500
Qy	4710		CCAGGCCGCGCCGACCGCCACAGCATCACCACCTCAACCAACCCCTTGACCTCAC	4769
Db	4501		CCAGGCCGCGCCGACCGCCACAGCATCACCACCTCAACCAACCCCTTGACCTCAC	4560
Qy	4770		AAGGAGCAGTGTCTGAGGCTGCACTGATGGCTCTCTCGGTGGACGCTCTCTGCTCCATC	4829
Db	4561		AAGGAGCAGTGTCTGAGGCTGCACTGATGGCTCTCTCGGTGGACGCTCTCTGCTCCATC	4620
Qy	4830		TGTGTGTCTTTGCCATGTCTTTGTGCCGGCAGCTTCACTCTGTCTCTCATTTAGGAG	4889

Db	4621	TGTTGGTCTTTTGCATGTCCTTTTGTGCCGGCCAGCTTCACTCTTTGTCCTCATTTGAGGAG	4681
Qy	4890	CGAGTCAACCGAGCCAAAGACACTGCAGCTCATTGGGGGGCCCTGTCCCCCACCCTCTACTCG	4949
Db	4681	CGAGTCAACCGAGCCAAAGACACTGCAGCTCATTGGGGGGCCCTGTCCCCCACCCTCTACTCG	4740
Qy	4950	CTTGGGCAACTTCTCTGGGACATGTGTAACTACTTGGTGCACGATGCATCGTGGTGCCTC	5009
Db	4741	CTTGGCAACTTCTCTGGGACATGTGTAACTACTTGGTGCACGATGCATCGTGGTGCCTC	4800
Qy	5010	ATCTTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCCCTGGCCAACTGCTCTCTCCCTG	5069
Db	4801	ATCTTTTCTGGCCCTTCAGCAGAGGGCATATGTGGCCCCCTGGCCAACTGCTCTCTCTCCCTG	4860
Qy	5070	CTGTGTCTACTACTGTATGCTGGTGCATCACACCCTCATGTACCAGGCTCCTTCTTCTC	5129
Db	4861	CTGTGTCTACTACTGTATGCTGGTGCATCACACCCTCATGTACCAGGCTCCTTCTTCTC	4920
Qy	5130	TTCTCCGTGCCAGCAGCAGCTATGTGGTGTCTCACCTGCATAAACCTCTTTATTGGCATC	5189
Db	4921	TTCTCCGTGCCAGCAGCAGCTATGTGGTGTCTCACCTGCATAAACCTCTTTATTGGCATC	4980
Qy	5190	AATGGAAGCATGCCCACCTTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG	5249
Db	4981	AATGGAAGCATGCCCACCTTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG	5040
Qy	5250	AGCCGATCTTGTAAACAGAGTCTTCTTATCTTCCCCCACTCTTCTGTGGGCCGGGGGCTC	5309
Db	5041	AGCCGATCTTGTAAACAGAGTCTTCTTATCTTCCCCCACTCTTCTGTGGGCCGGGGGCTC	5100
Qy	5310	ATTGACATGTCGGAAACAGAGGCATGGCTGTATGCTTTGAGCGCTTTGGAGACAGGCAG	5369
Db	5101	ATTGACATGTCGGAAACAGAGGCATGGCTGTATGCTTTGAGCGCTTTGGAGACAGGCAG	5160
Qy	5370	TTCCAGTCAACCTCGCTGGGAGGTGGTCGGCAAGAACCTCTTGTGCCATGTGTGATACAG	5429
Db	5161	TTCCAGTCAACCTCGCTGGGAGGTGGTCGGCAAGAACCTCTTGTGCCATGTGTGATACAG	5220
Qy	5430	GGGCCCCCTCTCTCTTTCACACTACTGTCTGCAGCAGCGAGCCAACTCTCTGCCACAG	5489
Db	5221	GGGCCCCCTCTCTCTTTCACACTACTGTCTGCAGCAGCGAGCCAACTCTCTGCCACAG	5380
Qy	5490	CCCAGGGTGAAGTCTCTGCCACTCTCTGGAGAGGAGCAGGATGTAGCCCGTGAACCG	5549
Db	5281	CCCAGGGTGAAGTCTCTGCCACTCTCTGGAGAGGAGCAGGATGTAGCCCGTGAACCG	5340
Qy	5550	GAGCGGGTGTCAAGAGGCCACCCAGGGGGATGTGTGTGGTGTGAGGAACTTTGACCAAG	5609
Db	5341	GAGCGGGTGTCAAGAGGCCACCCAGGGGGATGTGTGTGGTGTGAGGAACTTTGACCAAG	5400
Qy	5610	GTATACCGTGGGAGAGGATGCCAGCTGTGTGACCGCTTGTCTCTGGGGATTTCCCCCTGGT	5669
Db	5401	GTATACCGTGGGAGAGGATGCCAGCTGTGTGACCGCTTGTCTCTGGGGATTTCCCCCTGGT	5460
Qy	5670	GAG 5672	
Db	5461	GAG 5463	

Search completed: December 30, 2004, 13:41:53
Job time : 3248 secs

This Page Blank (uspto)

Result No.	Query \$			DB	ID	Description
	Score	Match	Length			
1	1648.8	24.2	6606	9	AY405471	Homo sapi
2	1628.8	23.9	6600	9	AY405473	Mus muscu
3	1559.4	22.9	6606	9	AY405472	Fan trogl
4	1150.8	16.9	4180	6	CB606356	AMGNNUC:S
5	904.4	13.3	4713	3	AK051920	Mus muscu
6	820.8	12.2	933	5	BQ959013	AGENCOURT
7	820.2	12.1	3759	3	CR627391	Homo sapi
8	805.8	11.8	3605	3	AK052916	Mus muscu
9	799	11.7	958	5	BQ955322	AGENCOURT
10	760.6	11.2	1934	3	BC036566	Mus muscu
11	751.8	11.0	914	5	BUS28031	AGENCOURT
12	748.2	11.0	969	5	BUR846021	AGENCOURT
13	733.2	10.8	884	4	BG775870	AGENCOURT
c 14	717.4	10.5	749	6	CD629699	AGENCOURT
15	704.8	10.4	730	6	CD629709	AGENCOURT
16	700.6	10.3	941	2	BE793677	AGENCOURT
17	684.4	10.1	880	4	BG674537	AGENCOURT
18	680.2	10.0	1415	3	BC038521	Mus muscu
c 19	678.2	10.0	705	6	CD629715	AGENCOURT
20	675.6	9.9	714	6	CD629703	AGENCOURT
c 21	673	9.9	681	6	CA314076	UI-CF-FNO
22	670.4	9.9	782	4	BG775975	AGENCOURT
c 23	655.4	9.6	707	6	CD629704	AGENCOURT
24	653.8	9.6	668	6	CD629707	AGENCOURT

Db 862 GACAACTCTACAACTCTTACTGCAATGATTGATGAAGAAATTTGGAGTCTAGTCTCTT 921
Qy 1047 TCCCGCTGTCTGGAGACGCTGAAGCCTTGATCTCTGGGAAGTACTTCTTGGACCA 1106
Db 922 TCCCGCATATCTGGAAGCTCTGAAGCGCTGCTGTTGGGAAGATCTCTGTATACACT 981
Qy 1107 GATACACTTTTACCCGGAAGCTATGCCAGGTGAACCGGACCTTCGAGGAGCTCAC 1166
Db 982 GACATCCAGCAACAGCGCAGGTATGGCTGAGGTGAACAGACCTTCCAGGAACCTGGCT 1041
Qy 1167 CTGCTGAGGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATG 1226
Db 1042 GTGTTCCATGATCTGGAAGCATGTGGAGAACTCAGCCCCAAGATCTGGACCTTCATG 1101
Qy 1227 AACGACAGTTTCAATGTGGCAATGCTCAGCGGCTCTCTGAGATGCAAGGA 1276
Db 1102 GAGAACAGCAAGAAATGGACCTTGTCCGATGCTGTTGGACAGCGGCAATGACCAC 1161
Qy 1277 -----TGAAGGAAGAGGAGCCAGCCAGACCTGGAGCGGGACCAATGG 1320
Db 1162 TTTTGGGAACAGCAGTTGGATGGCTTAGATGGACAGCCCAAGACATCTGGGGCTTTT 1221
Qy 1321 ----AGCCCTGCGATCTTCTGACCTCGGAGCGGTGCTACAGCTGGCAGAGCGCA 1376
Db 1222 GCCAAGCACCCAGAGGATGCCAGTCCAGTAATGGTTCTGTGTACACTGGAGAGACT 1281
Qy 1377 CAGCTGATGTGGGCACTGTGGGCACTGGGCGGCGAGTGCAGAGTGCCTGTCTTG 1436
Db 1282 TTCAACGAGACTAACAGGCAATCCGACCATATCTCGCTTCATGGAGTGTCTCAACTG 1341
Qy 1437 GACAGCTGAGGCGGACCTCAGAGCAGCCCTGGTGTGGCGGCTGCACTGCTC 1496
Db 1342 AACAGCTAGAACCCATAGCAACAGAAAGTCTGGCTCATCAACAAAGTCCATGGAGTCTG 1401
Qy 1497 GCGGAACATGATTTCTGGGCGGCGTCTCTTCTTGGGACCTGAGGACTCTTCAAGACCC 1556
Db 1402 GATGAGAGGAAGTTCTGGGCTGGTATTTGTTCACTGGAAT-----ACTCCAGGAGC 1455
Qy 1557 ACAGAGCACCCAGACCTGGGCGGCGGCGCACGTGGCGCATCAAAATCCGATGGAC 1616
Db 1456 ATTGAGC-----TGCCCCCATCATGTCAAGTACAAGATCCGAATGGAC 1497
Qy 1617 ATTGACGTGTACAGAGCAACATAAGATCAGGACAGAGTTTGGGACCTTGGGACCCAGCC 1676
Db 1498 ATTGACAAATGTGGAGAGCAATAAAATCAAGATGNNNNNNNNNNNNNNNNNNNNNN 1557
Qy 1677 GCGGACCCCTGACCGACCTGCGCTACGTGTGGGCGGCTTCTGTGTACCTGCAAGACCTG 1736
Db 1558 NNN 1617
Qy 1737 GTGAGCGTGCAGCGCTCCGCTGCTCAGCGGCGCCAAACCCCGGCGGCTCTACCTG 1796
Db 1618 NNN 1677
Qy 1797 CAGCAGATGCCCTATCCGTGCTATGTGAGACGTGTTCTGCGTGTGCTGAGCCGCTG 1856
Db 1678 NNN 1737
Qy 1857 CTGCGGCTCTTCTGACGCTGGCTGATCTACTCCGTGACACTCAGCAGTGAAGGCGGTG 1916
Db 1738 ATGCCCCCTTTCATGACGCTGGCTGGATTACTCAGTGGCTGTGATCATCAAGGGCATC 1797
Qy 1917 GTGCGGGAAGAGAGACGCGCTCGGGGACACCATGCGGCGCATGGGGCTCAGCGCGCG 1976
Db 1798 GTGTATGAGAGGAGGACCGCTGAAAGAGACCATGCGGATCATGGGCTTGGACAAAGC 1857
Qy 1977 GTGCTGTGGCTAGCTGTGTTCTCAGCTGCTCGGCGCTTCTGCTCAGCGCGCTG 2036
Db 1858 ATCTCTGGTTTACCTGTTTATAGTACCTTATCTCTTCTGTGAGCGCTGGGCTG 1917
Qy 2037 CTGTTCTGTGTGCTCAAGCTGGGAGACATCTCCCTTACAGCCACCGGCGGTGCTTTC 2096
Db 1918 CTAGTGTCTCATCTGAAGTTAGGAAACCTGTGTCCTTACAGTATCCAGCGTGTGTTT 1977

Qy 2097 CTGTTCTTGGACGCTTCCGGGTGGCCACGGTGACCCAGAGCTTCTGCTCAGGCGCTTC 2156
Db 1978 GTCTTCTGTCCGTGTTGCTGTGGTGACAATCTGCACTGCTTCTGATTTAGCACACT 2037
Qy 2157 TTCTCCGCGGCAAACTGGCTGGGCTCGGGGCTGCGGCTACTTCTCCCTTACTCTG 2216
Db 2038 TTCTCAGAGCCAACTGGCAGAGCTGTGGGGGATCATCTACTTACGCTGTACTCTG 2097
Qy 2217 CCCTAGCTGTGTGTGGCTTGGCGGACCGGCTGCCCGGGTGGCCCGCTGGCGCG 2276
Db 2098 CCCTAGCTGTGTGTGGCATGCGAGGACTACGTGGGCTTCACTCAAGATCTTCTGCT 2157
Qy 2277 AGCTGTGTGCGCGCTGGCTTCCGCTTCCGCTGCGAGAGCTTGGCTGTCTGGAGGAG 2336
Db 2158 AGCTGTGTCTCTGCTGGCTTTGGGCTTGGCTGTGAGTACTTTCCTCTTTTGGAGGAG 2217
Qy 2337 CAGGCGAGGCGCGCAGTGGCAACAGTGGGCAACCGGCT--ACGGCAGAGCTTCTC 2393
Db 2218 CAGGCAATGGAGTGCAGTGGGCAACCTGTTTGTAGAGTCTCTGTGGAGGAAGATGGCTTC 2277
Qy 2394 AGCTTGGCCAGGTCTCTGCGCTTCTGCTGTGGACCGGCGCTCTACGGCCTTGGCCACC 2453
Db 2278 AATCTCACCACTTTCGGTCTCCATGATGCTGTTTGACACCTTCTCTATGGGCTGATGACC 2337
Qy 2454 TGGTACTTGGAGCTGTGCTCCAGGCGCAGTACGGGATCCCTGAAACATGGAATTTTCT 2513
Db 2338 TGGTACTTGGAGCTGTCTTCCAGGCGCAGTACGGAAATTTCCAGGCGCTGATTTTCT 2397
Qy 2514 TTTTGGAGGAGTACTGTGTGCGGACCTCGGCCCCCAAGAGTCCAGCCCCCTTGGCCCCACC 2573
Db 2398 TGACCAAGTCTTACTGTTTGGGAGGAAGTGTAGAGAGGACCACTTGTGTTCCAAC 2457
Qy 2574 CCGCTG---GACCCAAAGTGTGTGTAGAGAGGCAACCGCCCGCTGAGTCTTGGCGTA 2630
Db 2458 CAGAAGAAATATCAGAAATCTGCATGGAGGAGGAACCCACCACTTGAAGCTTGGCGTG 2517
Qy 2631 TCCGTTCCGAGCTGTGGAGAGCGCTTCTGTTGGAAGCCCGCAGCCCTCGCGGGGCTC 2690
Db 2518 TCCATTGAGAACTGTGTAAAGTCTACCGAGATGGGATGAAGTGTGTCGATGGCGTG 2577
Qy 2691 AGCTGTGACTTCTACCAAGGCGCATACACCGCTTCTTGGGCCAACACGGGCGCGCAAG 2750
Db 2578 GCACTGAATTTTATGAGGCGCAGATCACTCTTCTTGGGCCACAAATGGAGCGGGAAG 2637
Qy 2751 ACCACCACTGTCTCATCTTGTAGTGGCTTCTCCACCCAGTGGTGGCTTCTGCTTTCATC 2810
Db 2638 ACACCACTGTCTCATCTGACCGGCTTGTCTCCCGCGACCTCGGCGCACCGCTACATC 2697
Qy 2811 CTGGGCCAGACCTCCGCTCCAGCATGGCGCCATCCGGCCCCACCTTGGGCGCTGTCTCT 2870
Db 2698 CTGGGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTTGGGGGCTGTGCC 2757
Qy 2871 CAGTACAACTGTCTGTTTGCATGCTGACCGTGGAGAGCAGCATCTGTTTCTATGGCGG 2930
Db 2758 CAGCATAACTGTCTGTTGACATGCTGACTGTGGAAGAACACATCTGTTTCTATGCCCGC 2817
Qy 2931 CTGAAAGGCTGTAGTGGCGCTGTAGTGGGCGCCAGCAGAGACCGTCTGTGCGAGGATGTG 2990
Db 2818 TTGAAAGGGCTCTCTGAGAAGCAGTGAAGCGGAGATGGAGCAGATGGCCCTGGATGTT 2877
Qy 2991 GGGCTGTCTC---CAAGCAGAGTGTGAGACTCGCCACCTCTCTGTGGGATGCAACGG 3047
Db 2878 GGTTCGCATCAAGCAGCTGAAAGCAAAAGCAAGCCAGCTGTGAGTGAATGAGAGA 2937
Qy 3048 AAGCTGTCCGTGGCCATTTGCTTTTGGGCGGCTCCCAAGTTGTATCTCTGAGAGGCT 3107
Db 2938 AAGCTATCTGTGGCTTGGCTTTTGGGCGGATCTAAGGTTGTCTTCTGGATGAACCC 2997
Qy 3108 ACAGCTGTGGATCTGCTTCCGCGCGGCTATTTGGAGAGTGTCTGCTCAATACCGA 3167
Db 2998 ACAGCTGTGTGACCTTACTCCCGCAGGGGAATATGGGAGCTGTGCTGTAATACCGA 3057

Db 842 AAGCGTGGACACCTTCTATGACAAATCTCAACTCTCTTATGCAATGATTTGATGAAGA 901
Qy 1027 CCTGGACAGCCACCCGCTGTCCCGCTGCTCTGGAGAGCGCTGAAGCCTCTGATCTCTCG 1086
Db 902 ACTTGGAGTCTAGTCTCTTCTCGAATATTTGGAGGCACTCAAGCACTGCTGTGTG 961
Qy 1087 GGAAGTACTCTTTGACACAGATACACCTTTTACCCGGAAGCTCATGCGCCAGGTGAACC 1146
Db 962 GAAAGATTCTCTATACACTGACACACAGCTTACAAGGCAGTCTATGCTGAGGTGAACA 1021
Qy 1147 GGACCTTCGAGAGCTACACCTGCTGAGGATGTCGGGAGGTGAGAGATGCTGGAC 1206
Db 1022 AGACCTTTCAGGAGTGTGCTGTGTTCCATGACCTGGAGGCACTGAGGAACTCAGCC 1081
Qy 1207 CCGGATCTTCACTTCAATGAACAGCACTTCAATGTGCCATGCTGAGCGGCTCT-- 1264
Db 1082 CCCAAATTTGACCTTCAATGAGAAACAGCAGAGATGACCTGTGCGGACGCTGTAG 1141
Qy 1265 -----GCGATGTCAGGATGAAGGAAGAGGAGGCCCA 1296
Db 1142 ACAGCAGAGGCAATGACCAGTTTGGGAAACAGAAGTTGGATGGATTAGATTGGACTGCC 1201
Qy 1297 GACCTGGAGGCGGGACACATGAGGCCCTGCGATCTCTTCTGACACCTGGAGCGGTG 1356
Db 1202 AAGACATCATGGCGTTTCTGGCCAGAAACCCAGAGATGTTTCAGTCCCCAAATGGCTCTG 1261
Qy 1357 GCTACAGCTGCGAGAGCCACACGCTGATGTGGGCACTGCTGGGCACTGCTGGGCGAG 1416
Db 1262 TGTATACCTGGAGAGAGCTTTCAATGAGAACCAACAGGCAATCCAGACGATATCTCGAT 1321
Qy 1417 TGACGGAGTGTCTTCTTGGACAGCTGGAGGGGCAACCTTCAGAGGAGCCCTGTTGT 1476
Db 1322 TCATGGAGTGTCAACTGTAACAGCTGGAACCCATTTCCGACAGAACTCAGGCTCATCA 1381
Qy 1477 CGGCGCCTCGAATGCTCGCGGAACATCGAATTTGCGGCCGCGTGTCTTCTTGGAC 1536
Db 1382 ACAAGTCCATGGAGCTGTGACGAGAGAAAGTCTTGGGCTGGCATCGTGTTCACAGG-- 1439
Qy 1537 CTGAGGACTCTTCAGACCCACAGAGCACCAACCCAGACCTGGGCCCGCCGACAGTGC 1596
Db 1440 -----CATCACTCAGATAGTGTGGAGCTGCCCATATGTAA 1477
Qy 1597 GCATCAAAATCCGATGACATGACGTGGTCAAGAGGACCAATAAGATCAGGACAGGT 1656
Db 1478 AGTCAAGATCCGGATGACATGACAACTGGAGAGAACTAAATAGATCAAGATGNN 1537
Qy 1657 TTTGGACCTTGGCCCGGACCCCTGACCGACCTCGGCTACCGACCTGGTGGGGGCT 1716
Db 1538 NNN 1597
Qy 1717 TCGTGTACCTGCAAGACCTGTGGAGCGTGCAGCGCTCGCGTGTCTCAGCGCGCAACC 1776
Db 1598 NNN 1657
Qy 1777 CCGGCGCGGCTCTACCTGCAGCAGATGSCCCTATCCGTGTATGTGACGACGTGTTCC 1836
Db 1658 NNN 1717
Qy 1837 TCGTGTCTCAGCGCGTGTGCGGCTCTCTCTGACGCTGCGCTGATCTACTCCGTGA 1896
Db 1718 TCGGGTCTATGAGCGGCTCAATGCGCCCTCTTCTATGACTCTAGCCTGGATCTACTCTG 1777
Qy 1897 CACTGACAGTCAAGCGCTGTGGGAGAGGAGACGCGCTGCGGACACCATGCGCG 1956
Db 1778 CTGTGATCATCAAGAGCATGTGTATGAGAGAGGCTCGGCTGAAGAGACCATGCGGA 1837
Qy 1957 CCATGGGCTCAGCGCGCGTGTCTGGCTAGGCTGGTCTCTCAGCTGCTCGGGGCT 2016
Db 1838 TCATGGGTCTGACAACTGATCTCTGCTTTAGTGTGTTGTAGCAGCTCATCCCTC 1897
Qy 2017 TCCTGCTCAGCGCGCGCTGTCTGTTTGTGCTCAAGCTGGGAGACATCTCCCTACA 2076
Db 1898 TGCTGTGAGCGCTGGCTGTGTTGTTGATCTTGAAGTTAGGAAACCTGTGCTGCCATA 1957

Qy 2077 GCCACCGGGCGTGTCTTCTGTTCTTGGCAGCCTTCGCGTGGCCACGGTGACCCAGA 2136
Db 1958 GTGACCCAGCGTGTGTCTTCTGTTCTGTTGTTGCCATGTTGACATCTCTACAGT 2017
Qy 2137 GCTTCTGCTCAGCGCCTTCTTCTCCGCGCAACCTGGCTCGGCTTCGCGCGCCTGG 2196
Db 2018 GCTTCTCATTAGCAGCGCTCTTCTCCGCTGCCAACCTGGCAGACGCTGTGGGGCATCA 2077
Qy 2197 CTTACTTCTCCTCTACTGCGCCTACGTGCTGTGTGTGTGGCTTGGCGGACCGGCTGCCG 2256
Db 2078 TCTACTTTCACGCTGTACTGCGCCTATGTGCTGTGCTAGCCTTGGCAGACTATGTGGCT 2137
Qy 2257 CGGCTGGCGCGTGGCGCGAGCTGCTGCTCGCGTGGCTTCTGGCTTCTGGCTGTGAGT 2197
Db 2138 TCTCCATCAAGATCTTGTGCTAGCTGCTCTCTGTTGGCTTTTGGATTTGGCTGTGAGT 2197
Qy 2317 GCCTGGCTCTGCTGGAGGAGCAGGCGGAGGCGCGCACTGTGCGACAA----CGTGGCACCC 2373
Db 2198 ATTTTCGCCCTTTTCGAGGAGCAAGGTATCGGGGTCCAATGGGACAACTCTCTTGGAGCC 2257
Qy 2374 GGCCTACGGCAGAGCTTCTCAGCCTGGCCAGGTCTCTGGCTTCTGCTGTGGAGCGG 2433
Db 2258 CGGTGAGGAGGACGGCTTCAATCTCACCACTGCAGTGTCTCATGATGCTCTTTCACCT 2317
Qy 2434 CGCTTACGGCTCGCCACCTGCTGTAAGCTGTGTCGCCAGGCGCAGTACGGGATCC 2493
Db 2318 TTCTATAGGCGTATGACATGTGATCTGAAAGCGCTTTTCCAGGACAGATATGGAATTC 2377
Qy 2494 CTGAACCATGAATTTTCTTTTCGAGGAGCTACTGTGCGGACCTCGGCCCCCAAGA 2553
Db 2378 CCAGGCCCTGTATTTTCTTTGTACCAAGTCTACTTGTGTTGGTGAGGAAATGTATGAGA 2437
Qy 2554 GTCCAGGCCCTTGGCCACCCCGCTG---GACCCAAAGTCTGTTGTAAGAGGACCGC 2610
Db 2438 AGAGCCACCTTGGTTCAGCCAGAGGAGGTGTGAGAAATCTGATGTAAGAGGAACCA 2497
Qy 2611 CCGGCTGAGTCTCGGCTGTCCGTTCGACCTTGGAGAGCGCTTCTCTGGAAGCCGCG 2670
Db 2498 CTCTATGAGGCTGGGGTGTCCATTCAGAACCTGTGTGAGGTTTACCGAGATGGCATGA 2557
Qy 2671 AGCCAGCCTCGCGGGCTCAGCTTGAATCTTACAGGCGCAATACCGCTTCTG 2730
Db 2558 AGTGTGTGTGGATGTGGCTTGGCGCTCAACTTTACGAAGGCGCAGATTAACCTCTCTCTG 2617
Qy 2731 GCCAAGCGGGCGGCAAGACACCACTGTCTCATCTTGAAGTGGCTCTTCCACCA 2790
Db 2618 GCCAATGGAGCGGGAAGACCAACCATGTCAATACTGACTGGGCTGTTTCCCCCAA 2677
Qy 2791 GTGTGTGCTCTGCTTCTCATCTTGGCCACGACGTCCGCTCCAGCATGGCCGCCATCCGCG 2850
Db 2678 CTTCTGGCACGGCTCATCTCTGGGAGGACATTCGCTCGAGATGAGCTCCATCCGCG 2737
Qy 2851 CCCACCTGGCGTCTGCTCCTCAGTACAAAGTGTCTTGTGACATGCTGACCGTGGACGAGC 2910
Db 2738 AGAACCTGGGAGTCTGTCCCAGCATAAATGTGCTGTTGACATGCTGACTGTGGAAGAGC 2797
Qy 2911 ACGTCTGTTCTATGGCGGCTGAAGGTCTGAGTGCCTGTAGTGGGCCCCCGAGG 2970
Db 2798 ACATCTGTTCTATGCGCGCTTAAAGGGCTCTCAGAGAAAGCACTGTAAGACAGAGATGG 2857
Qy 2971 ACCGTCTGCTGCGAGATGTGGGGCTGTCTC---CAAGCAGAGTGTGAGACTCGCCACC 3027
Db 2858 AGCAGATGGCCCTGATGTTGGCTTACCCCGAGCAAGCTGAAAGCAAAACGAGTCAAG 2917
Qy 3028 TCTCTGTGGATGCAACGGAAGCTGTCCGTTGCGCAATGCTTTGTGGCGGCTTCCCAAG 3087
Db 2918 TCTCAGGTGGATGACAGAAAGCTGTCTGTGGCTTGGGCTTCTGGGTGGATTCGAAG 2977
Qy 3088 TTGTTATCTCGGACGAGCTACGCTGCGTGGATCTGCTTCCCGCGCGCTATTTGG 3147
Db 2978 TTGTCATTTCTGGACGAGCCCAAGCCCGGGTGGAGCCCGTACTCTCGCAGGGGAATATGG 3037

QY 3148 AGCTGCTGCTCAATACCGAGAGGTGCGACGCTGATCCTCTCCACCACCACTTGGATG 3207
Db |||||
QY 3038 AACTCTCTGCTAAATACCGGCAAGGCGCCATATTTTGTCTACACCAATGAGCG 3097
Db |||||
QY 3208 AGGACAGAGCTGCTGGGAGACGCTGTGCTGTGGTGGGAGGTGGCGCTGTGCTGTGCTG 3267
Db |||||
QY 3098 AAGCTGACATCTTGGGAGACAGATTTGCCATCATTTCCCATGGGAAGCTGTGTGTGG 3157
Db |||||
QY 3268 GCTCCCACTCTTCTCGCGCGTCACTGCGGCTCGGCTACTACTGACGCTGTGTGAAG 3327
Db |||||
QY 3158 GCTCCTCCCTGTTTTTGAATAAACAGTTGGAAACGGGTACTATCTGACCCCTGTTAAGA 3217
Db |||||
QY 3328 CCGCCTGCCCCGTACACCAAT-----G 3351
Db |||||
QY 3218 AAGATGTGGAATCGTCCCTCAGTTCTCTGCAAAACAGTAGCAGCACCGTGTCTTGTCTGA 3277
Db |||||
QY 3352 AGAAGGCTGACACTGACATGGAGGGCAGTGTGGACACAGGAGGAAAGAAATGGCA 3411
Db |||||
QY 3278 AATAAGGAGGACAGTGTCTCAGAGCAGTTCTGATGTGCGCTGGGAGGACCATGAAA 3337
Db |||||
QY 3412 GCCAGGACAGAGTCCGCACTCTCAGCTGTGCGCCCTGGTACAGCACTGGGTGCCG 3471
Db |||||
QY 3338 GTGACACGCTGACCATCGATGTCTCTGCTATCTCCAACTCATCAGGAAGCACGTGTCTG 3397
Db |||||
QY 3472 GGGACCGCTGGTGGAGAGCTGCCACAGAGCTGGTGTGCTGCTGCCCTACACGGGTG 3531
Db |||||
QY 3398 AAGCCCGGCTGGTGGAGGACATTTGGGACAGAGCTGACCTATGTGCTGCCGTACGAAGCCG 3457
Db |||||
QY 3532 CCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTAGACACGCGGCTGGCGAGCTGA 3591
Db |||||
QY 3458 CGAAGGAGGAGCCTTTGTGGAATCTTCCATGATGATGACCGGCTCTCAGACCTGG 3517
Db |||||
QY 3592 GGCTCACTGCTACGGGATCTCGACACAGACCTCGAGGAGATCTTCTGAAAGTGTGG 3651
Db |||||
QY 3518 GCATCTCCAGTTATGGCATCTCGAGAGCACCCCTGGAAAGAAATATCTCAAAGTGGCTG 3577
Db |||||
QY 3652 AGGAGTGTCTGGGACACAGATATGAGAGTGGAGTGGACCTGGGGCAGCACTATGACAG 3711
Db |||||
QY 3578 AAGAGAGCGGGTGGATGCTGAGACCTCAGATGGTATCTTGGCAGCAAGACGAAACAGAC 3637
Db |||||
QY 3712 GCATTTGCTGCTAGACGTAACTCGGGCTCAAGATGCCCGCCACAGAGACAGCG---C 3768
Db |||||
QY 3638 GGGCTTCGGGGACAAGCAGAGCTGTCTGCACCCATTTACGGAAGATGATGCTTTGATC 3697
Db |||||
QY 3769 TGGAGAACGGGGAACACAGCTGGGTGACGCCAGAGACTGACAGGGCTCTGGGCGAGCG 3828
Db |||||
QY 3698 CCAATGACTCTGACATAGACCCAGAAATCCAGGAGACCGGACCTGCTCAGTGGGATGGACG 3757
Db |||||
QY 3829 CCGTGGG-----CCGGGTACAGGGCTGGGCACTGACCCGCCAGCAGCTCCAGGCCCTGC 3882
Db |||||
QY 3758 GCAAAGGCTCCTTACAGCTGAAGGGCTGGAAACTCACCCAGCAACAGTTTGTGGCCCTTT 3817
Db |||||
QY 3883 TTCTCAAGCGCTTTCTCTTGGCCCGCGCAGCCGCGGCTGTTCGCCCAAGATCGTGC 3942
Db |||||
QY 3818 TGTGGAAGAGGCTGCTGATTTGCCAGACGAGCCGGAAGGTTCTTTGCTCAGATTTGCTC 3877
Db |||||
QY 3943 TGCTGCTCCCTTTTGTGGGCTGGCCCTCGTGTTCAGCTCACTGTCGCTCCTTTTGGGC 4002
Db |||||
QY 3878 TGCAGCTGTCTTTGTTGATTTGCCCTGGCTCTTTCAGCTGATTTGTGCCACCCCTTTGGCA 3937
Db |||||
QY 4003 ACTACCCGGCTGCGGCTCAGTCCACACATGATACGGTGTCTCAGGTGTCTTCTCAGTG 4062
Db |||||
QY 3938 AGTACCCACGCTTGAACCTTACGCCCCCTGGATGATATATGAGCAGTATACATTTGTTCAGTA 3997
Db |||||
QY 4063 AGGACGCCCCAGGGGACCCCTGGAGCTGCCCGGCTGCTCGAGGGGCTCTGTCAGAGGAGCAG 4122
Db |||||
QY 3998 ATGATGCTCCCGAGGACATGGGACCCAGGAACTCTGATGCTCTGACCAAGATCCAG 4057
Db |||||
QY 4123 GACTGGAGGAGCCCGGAGTGCAGATAGCTCCACAGAGTTCTCGGACACAGAGATTCCTG 4182
Db |||||
QY 4058 GCTTTGGGACCCGCTGTATGAGAGGAAACCCAAATCCCAAGATACCCCTTGTGCTGGGG 4117
Db |||||
QY 4183 CTGAAG-----TGGCCAAGGTCTTGGCCAGTGGCA 4212

Db |||||
QY 4118 AGAGAGCTGGACCATCAGCCCCGTCCCCCAGAGCATCGTGGACCTCTTCCGAATGGAA 4177
Db |||||
QY 4213 ACTGGACCCAGAGTGTCTCCATCCCCAGCCTGCCAGTGTAGCCAGCCGGTGCCTGGCGGCC 4272
Db |||||
QY 4178 ACTGGACCATGAAGAACCCCTCACCTGCGTGGCGGCTAGCAGTGACAAATCAAGAAGA 4237
Db |||||
QY 4273 TGTGCCCCGACTCTCCCGGCTGACAGCTGGTGGTCCCCCTCCGCCCCAGGAGTGCACCGGCT 4332
Db |||||
QY 4238 TGTGCTGTGTGTCTCCCGCAGGGGCGAGGGGGCTGCGACCTCTCTCAGAGAAAAACAGAAA 4297
Db |||||
QY 4333 CTGGGAGAGTGGTTTCAGAACCTGACAGGCGGAGACCTGTCTGACTTCTCTCTCAAGACCT 4392
Db |||||
QY 4298 CCGACAGACATCTCTTCAAGATCTCAGAGGAAGAAATATTTTCAGATTAACCTGGTGAAGACGT 4357
Db |||||
QY 4393 ACCCGGCTCTGGTGGCGCCAGGGCTGAAAGACTAAGAAAGTGGGTGAATGAGGTCAAGTACG 4452
Db |||||
QY 4358 ACCTGACAGATCATAGCAAGAGCTTAAAGAAATGAAGATCTGGGTGAACGAGTTCGATG 4417
Db |||||
QY 4453 GAGGCTTCTGCTGGGGGGCCGAGA---CCAGGCTTGCCTCTGGGCGCAAGAGTGTGGGCC 4509
Db |||||
QY 4418 GCGGGTTTTCTTGGGTGTGAGTAATTTCTCAAGCACTTCTCCGAGCATGAAAGTTAATG 4477
Db |||||
QY 4510 GCTCAGTGGAGGAGTTGTGGCGCTGCTGAGTCCCTGCTGCGGGGCGCTCGACCGTG 4569
Db |||||
QY 4478 ATGCTATCAAGCAAAATGAAGAACTCTGAGCTGACCAAGGACAGCTCCGAGATCGCT 4537
Db |||||
QY 4570 TCCTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCT 4629
Db |||||
QY 4538 TCCTCAGCAGCCTGGGAAGGTTTCATGGCAGGGCTGGATACGAAAAACAATGTCAAGGTGT 4597
Db |||||
QY 4630 GGTTCAAACAAAGGCTGGCACTCCATGTTGGCTTTTGTCAACCGGAGCCAGCAACGCAA 4689
Db |||||
QY 4598 GGTTCAAATAACAAGGCTGGCATGCTATCAGTTTCTGTTCTGTAATGTCAACAANTGCCA 4657
Db |||||
QY 4690 TCCTCGTGTCTACCTGCCCCCAGGCGCGCCGCGCACGCCCCACAGCATCACCACACTCA 4749
Db |||||
QY 4658 TTCTCGGGCCAACTTCAGAGGGAGAGAACCCGAGCCAGTATGGAATCACTGCTTCA 4717
Db |||||
QY 4750 ACCACCCCTTGAACTCACCAAGGAGCAGCTGTCTGAGGCTGCACATGATGGCTCTCTCGG 4809
Db |||||
QY 4718 ACCACCCCTTGAACTCTCTATTAACAGCAGCTCTCAGAGGTGGCTCTGATGACCACTCTG 4777
Db |||||
QY 4810 TGACAGCTCTCTGCTCTCACTGTGTGTCTTTGCCATGTCTTTGTCGCCGCGCAGCTTCA 4869
Db |||||
QY 4778 TCAGCTCTCTGCTGTCTATCTGTGTCTCTTTCGATGTCTTGTCTCTGCTGCTGCTG 4837
Db |||||
QY 4870 CTCTGTCTCTCATTTGAGGAGCGAGTCAACCGAGCCAGCAAGCACTGACGCTCATGGGGGGCC 4929
Db |||||
QY 4838 TTGTGTTCTGATCCAGGAGCGTGTGAGCAAGCAAGCATCTTCAGTTCATCAGCGGCG 4897
Db |||||
QY 4930 TGTCCCCCACTCTACTGCTTGGCAACTTCTCTGGGACATGTGTAACTATTTGTTGTC 4989
Db |||||
QY 4898 TGAAGCTGTGCTACTTGGCTGTCAAATTTTGTCTGGGATATGTGCAATATATGTGTGCTC 4957
Db |||||
QY 4990 CAGCATGCTGCTGCTCATCTTCTGCGCTTCCAGCAGAGGGCATATGTGGCCCCCTG 5049
Db |||||
QY 4958 CTGCTACACTGGTCAATATCATCTTCTGCTTCCAGCAGAGTCTCTATGTGTCTCTA 5017
Db |||||
QY 5050 CCAACTGCTGCTCTCTGCTGCTACTGATGCTGCTGCTGATGCTGATCAGCAGCGCTCA 5109
Db |||||
QY 5018 CCAACTGCTGCTTCTAGCCCTCTGCTCTTGTGTATGAGGTGCTGATCACTCTCTCA 5077
Db |||||
QY 5110 TGTACCCAGCTCTTCTTCTTCTCCGTGCCAGCAGCAGCTATGTTGGTGTCACTCTGCA 5169
Db |||||
QY 5078 TGTATCCAGCATCTTGTGTCTCAAGATCCCGACCGGCTATGTGGTTCTCACCAGTG 5137
Db |||||
QY 5170 TAAACCTCTTATGGCATAATGGAAGCATGCGCACCTTTGCTGTGCTGCTCTCTCTG 5229
Db |||||
QY 5138 TGAACCTCTTCTATCGGATCAATGAGCAGTGTGGCCACTTTCGACTGAGCTGTGTACAA 5197
Db |||||
QY 5230 ATCAGAACTGCAAGGAGTGGAGCGGATCTTGAACAGAGTCTTCTCTTATCTTCCCCACT 5289
Db |||||

```
Db 5198 ACAATAAGCTCAATGACATCAATGACATCCTGAAAGTCTGTGTTTCTTATCTTTCCACATT 5257
Qy 5290 TCTGCTGGCGCGGGGCTCATTTGACATGCTGGTCCGAAACAGAGCCATGGCTGATGCTTTG 5349
Db 5258 TTTGCTGGGCGAGGGCTCATCGACATGTTGAAGAACAGGCCATGGCCGATGCCCTGG 5317
Qy 5350 AGCGCTGGGAGACAGGAGTTTCCAGTCAACCCCTGGCTGGAGGTGGTGGCAAGAAC 5409
Db 5318 AGAGTTTGGGAGAACCGCTTCGTCTCTCGCTCTCTGGGACTTGGTAGGACGGAACC 5377
Qy 5410 TCTTGGCCATGTTGATACAGGGCCCCCTCTCTCTCTTACACTACTGCTGACGACC 5469
Db 5378 TTTTGGCCATGCTGCGGAGGGGTGTTCTCTCTCACTACTTCTGATCCAGTACA 5437
Qy 5470 GAAGCCAACTCTCTGCACAGCCAGGGTGAGGTCTCTGCACTCTGGAGAGGAGGACG 5529
Db 5438 GATTTTTCATCAGGCCAGACCTGTAAAGCGAAGCTTCTCTTGAATGACGAGGATG 5497
Qy 5530 AGGATGTAGCCGTTGAAACGGGAGCGGGTGTCCAGGAGCCACCCAGGGGATGTGTTGG 5589
Db 5498 AGGATGTAGGGCGGAGAGACAGAGGATTTGACGGTGGCGGACAGAAATGACATCCTAG 5557
Qy 5590 TGCTCAGGAACTTGACCAAGGTATACCGTGGCAGAGGATGCCAGCTGTTGACCGCTGT 5649
Db 5558 AGATCAAGGAACCTGACCAAGATCTATAGGAGGAAGCGGAAGCCTGCAGCTGCACAGATCT 5617
Qy 5650 GCCTGGGATTCCTCGTGTAGTCTTTTGGTCTGCTGGTGTGAATGAGCAGGGAAGA 5709
Db 5618 GCATCGGATCCTCCGAGAGTCTTTGACTCTGGAGTTAACGGAGCTGGGAAGT 5677
Qy 5710 CGTCACAGTTTGGCATGTGAGCGGGGACACATTTGCCAGCAGGCGCGAGGCTGTGCTGG 5769
Db 5678 CAACAACCTTCAAGATGCTGACTGGAGACACCCCTGTGACACAGAGGGATGGTTCT-- 5735
Qy 5770 CAGGCCACAGCGTGGCCCGGAAACCCAGTGTGCGCACCTCAGCATGAGGATGCTGCCCTC 5829
Db 5736 ----TAAACAACATCTTATCAATATCCATGAATACACAGAAATGCGGCTACTGCCCTC 5791
Qy 5830 AATCGATGCCATCTTTCAGCTGTGACGGCGCGAGCACCTGGAGCTGTTGGCGCC 5889
Db 5792 AGTTTGACGCATCAGAGCTGTGACGGGAAGAGCATGTGGAGTCTTTGGCCCTCC 5851
Qy 5890 TGCGCGTGTCCCGAGGCGCCAGGTGTCGCCAGACCGCTGGCTCAGGCCCTGGCGGTCTGG 5949
Db 5852 TGAGGGAGTCCAGAAAGAGTGGCAGGTTGGTGAATGGCAATTCGCAACTGG 5911
Qy 5950 GACTCTCATGTGACGACCGCGCTGAGGCACCTACAGCGAGGGGAACAAACGCAAGC 6009
Db 5912 GCCTGGTAAAGTATGGAGAAAATATGCCAGTAACTACAGTGGCGGCAACAAACGAAAGC 5971
Qy 6010 TGGGACGGCCCTGGCGCTGTTGGGACCCAGCGCTGTTCTTGGACGAGCGGACCA 6069
Db 5972 TCTCAGAGCATGCTTGTATGTCGGGACCTCTCTGTTGTTCTGGATGAACCAACCA 6031
Qy 6070 CAGGCATGGACCCAGCGCGCGCTTCTCTTTGGAAACAGCTTTTGGCCGTGGTGGGG 6129
Db 6032 CAGGCATGGACCCTAAGCCCGGAGATCTTGTGGAATTTGGCCCTAGCATTTGTCAAG 6091
Qy 6130 AGGCGGTTTCAGTGTATCTCCATAGCATGGAGAGTGTGAAGCGCTCTGCTGCG 6189
Db 6092 AGGGAGATCTGTAGTCTTACATCTCATAGTATGGAAGATGTGAAGCTCTTTGTATCAA 6151
Qy 6190 GCCTAGCCATCATGTTGATGGCGGTTCCGCTGCTGGCGAGCCGCAACATCTCAAG 6249
Db 6152 GGATGGCCATAATGGTCAATGGAAAGTTTCAGGTGCTTTGGCAGTGTCCAAACATCTGAAA 6211
Qy 6250 GCAGATTCGCGCGGTCACACACTGACCTCTGCGGTGTCGCGCGCAAGGTCCC----- 6303
Db 6212 ACAGTTTGGAGATGTTTATACATAGTTGTACGATAGAGGCTCCAAACCTGACCTGA 6271
Qy 6304 AGCCGGCAGCGCCCTTCGTGGCGGCGGAGTTCCTCTGGGTGCGAGCTGCGGAGGCACATG 6363
Db 6272 AGCCTGTCCAGGAGTCTTTTGGACTTGGTTTCCGGGAAGTGTCTTAAAGAGAAACATC 6331
```

```
Qy 6364 GAGGCGGCTGCGCTTCCAGCTGCGCGGAGGCGCTGCGCCCTGCGCGCTCTTTG 6423
Db 6332 GAACATGCTTTCAGTACACAGCTTCCATCTCTTGTGAT--CTTACGAGGATATCA 6388
Qy 6424 GAGAGCTGGCGGTGACGCGCAGAGCAGCGGTGGAGGACTTTTCCGTGAGCCAGACA 6483
Db 6389 GCATCTCTCCAGAGCAAAAAGGAGCTCCACATAGAAAGACTACTCTCTCTCAGACAA 6448
Qy 6484 TGCTGGAGGAGTATCTTGTACTTCTCCAGGACCGGAGGAGACGAGGACCCGAG 6543
Db 6449 CACTTGACCAAGTATTTGTGAACCTTTCAGGACCAAGTATGATGACCATTTAAAG 6508
Qy 6544 AGCAG 6548
Db 6509 ACCTG 6513

RESULT 3
LOCUS AY405472
DEFINITION Pan troglodytes ABCA1 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY405472
VERSION AY405472.1 GI:39761446
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 606)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 606)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..606
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>606
/gene="ABCA1"
/locus_tag="HCM2235"

gene
Query Match 22.9%; Score 1559.4; DB 9; Length 606;
Best Local Similarity 52.7%; Pred. No. 5.2e-310;
Matches 3002; Conservative 0; Mismatches 2539; Indels 150; Gaps 14;

Qy 979 CCCTGCCAGACAGCAGCCTGAGCCCGCTGCTCGAGCTGATTTGGAGCCCTGGACAGCC 1038
Db 854 CTTNNNNNNNNNNNNNNNNNNCTCTTACTTAATGATTTGATGAAGAAATTTGAGTCTA 913
Qy 1039 ACCCGCTGTCCTGCTGAGAGCCCTGAGAGCTCTGATCTCTGGGAGCTACTCT 1098
Db 914 GTCTCTTTCCCGCATTTATCTGGAAGCTCTGGAAGCGCTCTGTTGGGAAGATCTCTGT 973
Qy 1099 TTGACACAGATACACCTTTTACCCGGAAGCTCTATGGCCAGGTGAAACCGGACCTTCGAGG 1158
```

Db 974 ATACACCTGACATCCAGGCACAAAGGACGCTCATGGCTGAGGTGAACAAGACCTTCACAGG 1033
Qy 1159 AGCTACCTGCTGAGGGATGTCCGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCA 1218
Db 1034 AACTGGCTGTGTTCCATGATCTGGAAGGCATGTGGGAGGAACCTCAGCCCAAGATCTGGA 1093
Qy 1219 CTTTCATGAACGACAGTTCCAAATGTGGCCATGCTGCAGCGGCTCTCCAGATGCAGGA-- 1276
Db 1094 CTTTCATGGGAACAGCAAGAATAGAACCTTGTCCGATGCTGTTGGACAGCAGGGACA 1153
Qy 1277 -----TGAAGGAAGAGCAGCCACAGCTCTGGAGCCGGGA 1312
Db 1154 ATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGACAGCCCAAGACATCGTGG 1213
Qy 1313 CCACATGG----AGGCCCTGGGATCCCTTTCTGGACCTCTGGAGCGGTGGCTACAGCTGGC 1368
Db 1214 CGTTTGTGGCAAGCACCCAGAGGATGTCCAGTCCAGTAATAGTGTCTGTGTACACCTGGA 1273
Qy 1369 AGGACGCACACGCTGATGTGGGGCACCTGTGTGGSCACGCTGGGCCGAGTGACGAGTGCC 1428
Db 1274 GAGAAGCTTTCAACGAGACTAAACAGGCAATCCGGACCAATATCTCGCTTCATGGAGTGTG 1333
Qy 1429 TGTCTTTGGACAAGCTGGAGCGGCACCTCAGAGGCAGCCCTGTGTGCGGGGCCCTGC 1488
Db 1334 TCAACCTGAACAAGCTAGAACCCATAGCAACAGAGTCTGGCTCATCAACAAGTCCATGG 1393
Qy 1489 AACTGCTCGCGGAACATCGAATTTCTGGCCGGCGTGTCTTCTTGGGACCTGAGGACTCTT 1548
Db 1394 AGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTATGTGTCTCACTGGAAATT-----ACTC 1447
Qy 1549 CAGACCCACAGAGCACCCAAACCCAGACCTGGGCCCGGCCAGCTGGCGCATCAAAATCC 1608
Db 1448 CAGGACGATTGAGC-----TGCCCCATCATGTCAAGTACAAGATCC 1489
Qy 1609 GCATGGACATTGACGTGTGTCAGAGGACCAATAGATCAGGGAACAGCTTTTGGACCCCTG 1668
Db 1490 GAATGGACATTGACAAATGTGGAGAGGACAAATAAAATCAAGGATGNNNNNNNNNNNN 1549
Qy 1669 GCCACGCGGGACCCCTGACCGACTGCGCTACGTGTGGGGGGGCTTCGTGTACTGTC 1728
Db 1550 NNN 1609
Qy 1729 AAGACCTGGTGGAGCGTGCAGCCGTCGCTGCAGCGGCCCAACCCCGCGCGGCC 1788
Db 1610 NNN 1669
Qy 1789 TCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGGACGACGTGTTCTCGCTGTGCTGA 1848
Db 1670 NNN 1729
Qy 1849 GCGGTCGCTGCGCTCTTCTGACGCTGGCTGGATCTACTCGGTGACACTGACAGTGA 1908
Db 1730 GCGGTCATGCCCCCTCTTCATGACGCTGGCTGGATTTACTCAGTGGCTGTGATCATCA 1789
Qy 1909 AGGCCGTGTGGGAGAGAGACGCGCTGCGGGACACCATGCGCGCCATGGGGCTCA 1968
Db 1790 AGGCGATCGTGTATGAGAAGAGGACACGGCTGAAAGAGACCATCGGATCATGGGCGCTGG 1849
Qy 1969 GCGCGCGGTCCTGCGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTGCTCAGCG 2028
Db 1850 ACAACAGCATCCTCTGGTTTAGCTGGTTTCAATTAGTAGCCTCATTTCTTCTTGTGAGCG 1909
Qy 2029 CCGCGCTGCTGGTTCCTGGTGTCAAGCTGGGGGACATCTCCCTCAAGCCACCCGGGCG 2088
Db 1910 CTGGCTGCTGGTGTGATCTCTGAAGTTAGGAACCTGCTGCCCCCTACAGTGATCCACAGCG 1969
Qy 2089 TGGTCTTCTGCTTCTGGCAGCTTCGCGGTGGCCAGGTGACCCAGAGCTTCTGCTCA 2148
Db 1970 TGGTGTGTGCTTCTGCTGCGGTGTTGTGTGGTGAACAATCTCGAGTGTCTCTGATTA 2029
Qy 2149 GCGCTTCTTCTCCCGGCGCAACCTGCGCTGCGGCTGCGGCGGCTGCGCTACTTCTCCC 2208
Db 2030 GCACACTCTTCTCAGAGCCAACTGGCAGCAGCTGTGGGGGCATCATCTACTTCAAGC 2089

Qy 2209 TCTAACCCTGCCCTACGTGCTGTGTGGCTTGGCGGACCGGCTGCCGCGGGTGGCGCG 2268
Db 2090 TGTACCTTACCCTACGTCCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCAAGA 2149
Qy 2269 TGGCCGCGAGCCCTGCTGTGTGCCCGTGGCCCTTCCGGCTTCGGCTGCGAGAGCTGGCTGCG 2328
Db 2150 TCTTCGCTAGCTGCTGTCTCTCTGTGGCTTTTGGTTTGGCTGTGAGTACTTTTGGCCCTTT 2209
Qy 2329 TGAAGAGCAGGGCGAGGGCGCGCAGTGGCAACAAGTGGGCAACCCGGCCT---ACGGCAG 2385
Db 2210 TTGAGGAGCAGGGCATTTGAGTGCAGTGGGACAACCTGTTTGAGAGTCTCTGTGTGGAGGAAG 2269
Qy 2386 ACCTCTTACGCTGGCCAGCTCTCTGGCCCTTCTGTCTGTGGACGCGGCGCTCTACGGCC 2445
Db 2270 ATGGCTTCAATCTCAACACTTCCTGGTCTCAATGATGCTGTTTGACACCTTCTCTATGGGG 2329
Qy 2446 TCGCCACCTGGTACTCTGGAAGCTGTGTGCCAGGCCAGTACGGGATCCCTGAAACCATGGA 2505
Db 2330 TGATGACCTGTATCATTTGAGGCTGTCTTCCAGGCCAGTACGGAATTTCCAGGCCCTGT 2389
Qy 2506 ATTTTCTTTTCGGAGAGCTACTGTGGGGAACCTCGGCCCCCAAGAGTTCAGACCCCTT 2565
Db 2390 ATTTTCTTGCACCAAGTCTCTACTGTGTTGGTGAGGAAGTGTGAGAGGACACCCCTG 2449
Qy 2566 GCCCAACCC---CGCTGGACCCAAAGGTGCTGTAGAGGACACCGCCCGGCTGAGTC 2622
Db 2450 GTTCCAGCCAAAAGAGATGTCAAAATCTGCAATGGAGGAGAACCCACCACCTGAAGC 2509
Qy 2623 CTGGCGTATCCGTTCCAGCTCGAGAAGCGCTTCTCTGGAAGCCCGCAGCCAGCCCTGC 2682
Db 2510 TGGCGTGTCCATTCAGAACCTGTGTAAGTCTACCGAGATGGATGAAGTGGCTGTNN 2569
Qy 2683 GGGGCTCAGCTTGGACTTCTACAGGGCCACATCAACCGCTTCTTGGGGCCAAACGGGG 2742
Db 2570 ATGGCTTGGCACTGAAATTTTATGAGGGCCAGATCACCTCTTCTCTGGGCCACAATGGAG 2629
Qy 2743 CCGCAAGACCAACCCCTGCTCATCTGTAGTGGCTCTTCCACCCAGTGGTGGCTCTG 2802
Db 2630 CGGGAAGACGACANNNGTCAATCTGACCGGGTGTTCCTCCCGACCTCGGGGACCG 2689
Qy 2803 CTTTCATCTTGGGCCACGACGCTCCGCTCCAGCATGCGCGCCCATCCGGCCCCACCTGGGCG 2862
Db 2690 CTTACATCTTGGGAAGAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGG 2749
Qy 2863 TCTGCTCTAGTACAACGCTGCTTTGACATGCTGACCGTGGAGAGCAGTCTGTGTTCT 2922
Db 2750 TCTGTCCCAGCATAAACGCTGCTTTGACNNNGTGACTGTGGAAGAACACATCTGTGTTCT 2809
Qy 2923 ATGGCGGCTGAAGGCTCTGAGTGGCGCTGTAGTGGGCCCGCCGAGCAGGACCGTCTGCTGC 2982
Db 2810 ATGCCGCTTGAAGGGCTCTCTGAGAAGCAGCTGAAGCGGAGATGAGACAGATGCC 2869
Qy 2983 AGGATGTGGGGCTGGTCTC---CAAGCAGAGTGTGCAGACTCGCCACCTCTCTGGTGGGA 3039
Db 2870 TGAATGTGTTTGCCATCAAGCAAGCTGAAAGCAAAACAAGCCAGCTGTCAAGTGGAA 2929
Qy 3040 TGCAACGGAAGCTGTGCTGGGCCATTTGCTTTTGTGGGCGGCTCCCAAGTTGTTATCCTGG 3099
Db 2930 TGCAGAGAAAGCTATCTGTGGCTTGTGGCTTTGTGCGGGGATCTAAGTTGTCTTCTGG 2989
Qy 3100 ACAGACTACGGCTGGGTGGATCTGCTTCCCGCCGCGGTATTTGGAGCTGCTGCTCA 3159
Db 2990 ATGAACCCACAGCTGGTGTGGACCCCTTACTCCCGAGGGGAATATGGAGCTGCTGCTGA 3049
Qy 3160 AATACCGAGAAGTTCGCACGCTGATCTCTCCACCCACCACTGGATGAGGACAGTGC 3219
Db 3050 AATACCGACAGGCCGACCATTTATCTCTACACACACATGATGAGTGAAGCGACATCC 3109
Qy 3220 TGGGAGACCGTGTGCTGTGGTGGCAGGTGGCGCTGTGTGCTGTGGTCTCCCACTCT 3279
Db 3110 TGGGGGACAGGATGGCCATCATCTCCATGGGAAGCTGTGCTGTGGGGCTCTCTCCCTGT 3169

QY 3769 TGGAGAACGGGGAACAGCTGGGTACGGTCCAGAGACTGACCGAGGGCTCTGGGGCAGACG 3828
Db 4199 CCAATGACTCTGACATAGACCCAGAAATCCAGGGAGACCGACCTGCTCAGTGGGATGGACG 4258
QY 3829 CCGTGGG-----CCGGGTACAGGGCTGGGCACTGACCGCCAGCAGCTCCAGGCGCTGC 3882
Db 4259 GCAAAGGCTCCTACAGCTGAGAGGGCTGGAACTCACCCAGCAACAGTTTGTGGCCCTTT 4318
QY 3883 TTCTCAAGCGCTTTCTGCTTTGCCGCGCAGCGCCGCGCTGTTTGGCCAGATCGTGC 3942
Db 4319 TGTGGAAGAGGCTGCTGATTTGCCAGAGCGGAGCGGAAGGGTTCTTTGCTCAGATTGTCC 4378
QY 3943 TGCCTGCCCTCTTTGTGGGCTGGCCCTGCTGTTTCAAGCTCATCGTGCCTCTTTGGGC 4002
Db 4379 TGCAGCTGTCTTTGTTGATTTGCCCTGCTTTCAGCTGATTTGTCACCCCTTTGGCA 4438
QY 4003 ACTACCGGGCTCTGGGCTCAGTCCACACCATGATGAGGTGCTCAGGTGCTCTTCAGTG 4062
Db 4439 AGTACCCCAAGCTTGAACCTCAGCCCTGGATGTATATAGACAGTATACATTTGTCAATA 4498
QY 4063 AGGACGCCCGCAGGGGACCTTGGACGTGCCCGGCTGCTCGAGGGGCTGCTCAGGAGGCGAG 4122
Db 4499 ATGATGCTCCGAGGACATGGCCACCCAGGAACTCTCTGAATGCTCTGACCAAGATCCAG 4558
QY 4123 GACTGGAG 4130
Db 4559 GCTTTGGG 4566

RESULT 6

BQ959013
LOCUS AGENCOURT 8958795 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460431
5', mRNA sequence.

ACCESSION BQ959013
VERSION BQ959013.1 GI:22374491
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LNCM2647 row: b column: 16
High quality sequence stop: 677.
Location/Qualifiers

FEATURES

source

1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6460431"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN	Query Match	Best Local Similarity	Score	DB 5;	Length	933;
	Matches	884;	Conservative	0;	Mismatches	33;
					Indels	6;
					Gaps	3;
QY	2982	CAGGATGTGGGGTGTCTCTCCAAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATG	3041			
Db	1	CAGGATGTGGGGTGTCTCTCCAAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATG	60			
QY	3042	CAACGAAGCTGTCCTGGCCATTGCTTTGGGGGGTCTCCCAAGTGTGTTATCTCTGGAC	3101			
Db	61	CAACGAAGCTGTCCTGGCCATTGCTTTGGGGGGTCTCCCAAGTGTGTTATCTCTGGAC	120			
QY	3102	GAGCCTACCGCTGGCTGGATCCTGTTCCCGCCGGGTATTTGGGAGTGTCTGCTCAAA	3161			
Db	121	GAGCCTACCGCTGGCTGGATCCTGTTCCCGCCGGGTATTTGGGAGTGTCTGCTCAAA	180			
QY	3162	TACCGAAGGTGCGACGCTGATCTCTCCACCCACACCTGATGAGGAGGAGCTGCTG	3221			
Db	181	TACCGAAGGTGCGACGCTGATCTCTCCACCCACACCTGATGAGGAGGAGCTGCTG	240			
QY	3222	GGAGACCGTGTGCTGTGGTGGCAGGTGGCCGCTTGTGCTGTGCTGCCACTCTTC	3281			
Db	241	GGAGACCGTGTGCTGTGGTGGCAGGTGGCCGCTTGTGCTGTGCTGCCACTCTTC	300			
QY	3282	CTGCGCCGTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCGCTGCCCCCTG	3341			
Db	301	CTGCGCCGTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCGCTGCCCCCTG	360			
QY	3342	ACCACCAATGAGAAGCTGACACTGACATGGAGGGCAGTGTGCACACCGAGGAGAAAG	3401			
Db	361	ACCACCAATGAGAAGCTGACACTGACATGGAGGGCAGTGTGCACACCGAGGAGAAAG	420			
QY	3402	AAGAATGGCAGCCAGCGCAGCAGAGTCGGCACTCTCAGCTGTGGCCCTTGGTACAGCAC	3461			
Db	421	AAGAATGGCAGCCAGCGCAGCAGAGTCGGCACTCTCAGCTGTGGCCCTTGGTACAGCAC	480			
QY	3462	TGGGTGCCCGGGGCA CGGCTGTGGAGGAGCTGCCACACAGCTGTGCTGTGCTGCC	3521			
Db	481	TGGGTGCCCGGGGCA CGGCTGTGGAGGAGCTGCCACACAGCTGTGCTGTGCTGCC	540			
QY	3522	TACACGGGTGCCCATGACGCGAGCTTCCGACACTCTTCCGAGAGCTAGACACGCGGCTG	3581			
Db	541	TACACGGGTGCCCATGACGCGAGCTTCCGACACTCTTCCGAGAGCTAGACACGCGGCTG	600			
QY	3582	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCTCGAGGAGATCTTCTG	3641			
Db	601	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCTCGAGGAGATCTTCTG	660			
QY	3642	AAGTGTGGAGAGTGTGCTGGGACACAGATATGAGAGATGCGAGCTCGGGGAGCAC	3701			
Db	661	AAGTGTGGAGAGTGTGCTGGGACACAGATATGAGAGATGCGAGCTCGGGGAGCAC	720			
QY	3702	CTATGCACAGGCAATGCTGGCTTAGAGTAACCTCGGCTCAAGATGCGCCACAGAG	3761			
Db	721	CTATGCACAGGCAATGCTGGCTTAGAGTAACCTCGGCTCAAGATGCGCCACAGAG	780			
QY	3762	ACAGC--GCTGGAGAACGGGG--AAACAGAGTGGGTTCAGCCCCAGAGACTGACAGGGCTCT	3818			
Db	781	ACAGCAGCTGGAGAACGGGGGNAACCCAGCTGGGTTCAGCCCCAGAGACTGACAGGGCTCT	840			
QY	3819	GGGCGACAGCGCCCTGGG---CCGGGTACAGGGCTGGGCACTGACCCGCGAGCTCCAG	3875			
Db	841	GGGCGCAAAACCCCGGGGGCCCGGGTACAGGGCTGGGCACTGCAACCCCTTCACTTCG	900			
QY	3876	GCCCTGCTTCTCAAGCGCTTTCT	3898			
Db	901	GCCCCCGCTTTTCAAGGGCTTTT	923			

RESULT 7

CR627391
LOCUS CR627391 3759 bp mRNA linear HTC 03-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp781N1972 (from clone DKFZp781N1972).
ACCESSION CR627391
VERSION CR627391.1 GI:50949875
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3759)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY

COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp781N1972) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp781N1972
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

1..3759
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781N1972"
/tissue_type="retina"
/clone_lib="781 (synonym: hicc4). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ATP-binding cassette, sub-family A member 4,
differentially spliced"

gene

/gene="DKFZp781N1972"
141..3338
/gene="DKFZp781N1972"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10486.1"
/db_xref="GI:50949876"

/translation="MDVVVLRHVPPEAKLVEICQELIFLLPNKFKRAYASLFRELE
TLADLGLSSFGISDTPLEIEFLKVTEDSDGLPAGGAQKRENVNPRHPCLGPREKA
GQTQDSNVCSGAPAAHPEGOPPEPEPCGPOLNTGTLVQLVHQVALLVKRFOHTIR
SHKPLAQIVLPATFVFLALMSLIVIPFGYPALTLHPWYQGYTFPSNDEPGSEQ
FTVLADVLNPKFGNRCLEKQWLPPEYPCGNSSTPWSFNITQFQKQKQVQNP
SPSCRSSTREKLMLEPEGAGLPPQRTORSTELQDLTDRIISDFLKYTPALI
RSSLSKFWNEKYGIGSLKFPVPTITGEALVGLSDIGRMVNSGGPITREASK
BIPDLKHELEDNIKFNKNGHWALVFLNVAHNAIIRASLPKDSPEYKIGITVIS
QPLNLTKEOLSEIPLVTSVDVAIVCI FMSFVPSFVLVLIQERVNKSKHLOFIS
GVSTTYVWTFPLDMNYSVAGLVGIFIGFOKAVTSPENIPALVALLLYGMV
IPMYIPASFLFDEPSTVALSCANLFGINSALTFILEFPENRTLLRNVALRKL
LIVFPFLCHGLGDLALUSQAVTDVYARFGEHSANPHWDLIGKUNLFAMVEGVVVF
LLTLVQHFGLSLQWIAIEPTEPIVDEDDVAERQRIITGNGKTDILRLHELTKYIP
TSSPVDRLCVGPGRGFLGLVSGAGKTTFKMLTGDVTVTSGDATVAGKSLTN
ISEVHONGYCPQDAIDELLTGREHLYLARLGPVAAEIEKVANNYSIKSLGTVVA
DCLAGTSGGNKRKLSTALIGCPPLVLLDEPTGMDPOARRMLNVIIVISREGRA
VLNLSHSEECALCTRLAIWVGAFCRMGTIOHLKRFQDGYIVTWMKIKSPKDLIP
DLNPVEQFQNFPGSVQRHRYNMLQFVSSSLARIFQLLSHKOSLLIEEYSVTP
TTLDQVFVNFKAQQTESHDLPLHRAAGASRAQAD"

ORIGIN

Query Match 12.1%; Score 820.2; DB 3; Length 3759;
Best Local Similarity 55.7%; Pred. No. 7.8e-158;
Matches 1747; Conservative 0; Mismatches 1338; Indels 54; Gaps 7;

3437 TCAGCTGCTGCGCCCTGGTGTACACACTGGGTGCCCCGGGCAACGGCTGGTGGAGAGCTGCC 3496
134 TGAGCTGATGATGTAGTTCTCCACCATGTTCCAGAGGCAAGCTGGTGGATGCTGG 193
3497 ACACAGAGCTGTGTGTGCTGCCCTACACGGGTGCCCATGACGGAGCTTCCGCACACT 3556
194 TCAAGAACTTATCTCTCTTCCAAATAAGAACTTCAAGCACAGACATATGCCAGCCT 253
3557 CTTCCGAGAGCTAGACACCGCGCTGGCGAGCTGAGGCTCACTGCTAGGGATCTCCGA 3616
254 TTTCCAGAGAGCTGGAGGAGACCTGGCGGACCTTGGTCTCAGACATTTTGGAAATTTCTGA 313
3617 CACCAGCCTCGAGGAGATCTTCTGAAAGTGTGTGAGGAGTGTGTGCGGACACAGATAT 3676
314 CACTCCCTCGAAGAGATTTTCTGAAAGTCAACGGAGGATCTGATTGAGGACCTCTGTT 373
3677 GGAGATGCG-----AGCTGGGCGACGACCTATGACAGGCATTTGCTGGCTAGAGCT 3730
374 TCGGGTGGCGCTCAGCAGAAAGAGAAAGCTCAACCCCGACACACCTCTGCTGGGTCC 433
3731 AACCTCGGGCTCAAGATGCCCGCACAGAGACAGCGCTGGAGAACGGGAACAGCTGG 3790
434 CAGAGAGAGCTGGAAGACACACCCAGGACTTCAATGTCTGCTCCAGGGGCGCCGGC 493
3791 GTCAGCCCCAGAGACTGACAGGGCTCTGGGCGCAGACCGCGTGGCGGTACAGGGCTG 3850
494 TGCTCACCACAGAGGGCGAGCTCCCGCAGAGCCAGAGTGCCAGGCGCCGAGCTCAACAC 553
3851 GG-----CACTGACCCCGCAGAGCTCCAGGCGCTGCTTCTCAAGCGCTTTCTGCTTGC 3904
554 GGGGACACAGCTGTCTCCAGCATGTGCGAGCGCTGCTGGTCAAGAGATTTCAACACAC 613
3905 CCGCGCAGCGCGCGCGCTTTCGCCAGATGCTGCTGCTGCTGCTGCTTCTTGTGGCGCT 3964
614 CATCCGAGCCACAGGACTTCTTGGCGCAGATGCTGCTCCCGGCTACCTTGTGTGTTT 673
3965 GGCCTCTGCTTACGCTCATCTGCTGCTCTTTTCGGGCACTACCGGCTCTGCGGCTCAG 4024
674 GGCTCTGATGCTTCTTATTGTTATCCCTCTCTTTTGGCGAATACCCGCTTGAACCTTCA 733
4025 TCCACCATGTACGGTGTCTAGGTGCTCTTCTTCAAGTAGAGAGCGCCCGGAGACCTGG 4084
734 CCGCTGGATATATGGCGAGCAGTACACCTTCTTCAAGATGAGTAAACAGGCGAGTGA 793
4085 ACGTGCCCGCTGCTCGAGGCGCTGCTGAGGAGCAGAGCTGAGGAGCCCGCAGTGA 4144
794 GTTCAGGCTACTTGACAGCTGCTCTCTGAAATAGCCAGGCTTTGGCAACCGCTGCTGAA 853
4145 GCA-----TAGCTCCCAACAGGTTCTCGGCACCAAGAGT 4177
854 GGAAGGTGGCTTCGGAGTACCCCTGTGGCAACTCAACACCTCGAAGACTCTCTTCTGT 913
4178 TCCTGTGAAGTGGCAAGGTCTTGGCCAGTGGCAACTTGGACCCCGAGAGTCTCCATCCCC 4237
914 GTCCCCAACAATCAACCCAGCTGTTCCAGAGCAGAGAAATGACAGAGTCAACCCCTTACC 973
4238 AGCTGCCAGTGTAGCAGCCCGCTGCTGCGCGCTGCTGCGGCTGCTGCGGCTGCGAGC 4297
974 ATCTGCAAGTGTAGCAGCAGGAGAGCTCACCATGCTGCGAGAGTCCCGCGAGGTGC 1033
4298 TGGTGTCTCCCTCCGCCCCCAGGAGTGAACCGGCTCTGCGGGAAGTGGTTCAAGAACCTGAC 4357
1034 CGGGGGCTCCCGCCCCCAGAGAAACACAGCGCAGACGCGAAATTTCTACAAGACTGAC 1093
4358 AGCGCGGAACCTGTCTGCTTCTTGGTCAAGACCTACCGCGCTGCTGCTGCGGAGGCT 4417
1094 GGACAGGAACATCTCCGACTTCTTGGTAAAAACGCTATCTCTGCTTATAAAGACAGCTT 1153
4418 GAAGACTAAGAGTGGGTGAATAGGTTCAGGTACGGAGGCTTCTCGCTGGGGGGCGAGA 4477
1154 AAAGAGAAATTCGGGTCAATGAACAGAGGTATGAGGAATTTCCATTGGAGGAAGCT 1213
4478 CCCAGGCTGCC---CTCGGGCCAAGAGTTGGGCGCTCAGTGGAGGAGTGTGTGGCGCT 4534

Db 1214 CCGAGTCGTCCTCCCATACGGGGAGACACTTGTGGGTTTTTAAGCGACCTTGGCCCGGAT 1273
Qy 4535 GCTGAGTCCCTGCTCGCGGGGGCCCTCGACCGTGTCTCGAATAACCTCAAGCCTGGGC 4594
Db 1274 CATGAATGTGAGCGGGGCCCTATCACTAGAGAGCCCTCTAAGAAATACCTGATTTCTT 1333
Qy 4595 TCACAGCCTGATGCTCAGGACAGTCTCAAGATCTGTTTCAACAACAAAGGCTGGCACTC 4654
Db 1334 TAAACATCTAAGAACTGAAGACAACTAAGGTGTGTTTAATAACAAGGCTGGCATGC 1393
Qy 4655 CATGGTGGCTTTGCTCAACGAGCCAGCAAGCAATCCTCGTGCTCACTGCCCCCAGG 4714
Db 1394 CTTGGTCAGCTTTCTCAATGTGGCCCAACAGCCCATCTTACGGGCCAGCCCTGCCCCAAGGA 1453
Qy 4715 CCGGGCCCGGCACCCACAGCACTCAACCACTCAACACCCCTTGAACCTCAACCAAGGA 4774
Db 1454 CAGGAGCCCGAGGAGTATGGAATCACCGTCAATAGCCAAACCCCTGAACCTGACCAAGGA 1513
Qy 4775 GCAGCTCTGAGGCTGCACTGATGGCTCTCTGGTGGAAGTCTCTGCTCTCCATCTGTGT 4834
Db 1514 GCAGCTCTGAGATTAACAGTGTGACCACTTCACTGATGATGCTGTGGTTGCCATCTCGT 1573
Qy 4835 GGTCTTTGCCATGCTCTTGTCCGGCCAGCTTCACTTCTGTCTCTCAATGAGGAGCGAGT 4894
Db 1574 GATTTTCTCCATGCTCTTGTCCCGCAGCAGCTTTGTCTTTATTTGATCCAGGAGCGGGT 1633
Qy 4895 CACCGAGCCAAAGCACCTGAGCTCATGGGGGGCTGTGCCCAACCCCTCTACTGGCTTGG 4954
Db 1634 GAACAAATCCAGACCTCCAGTTTATCAGTGGAGTGAGCCCCAACCACTACTGGGTGAC 1693
Qy 4955 CAATTTCTCTGGGAATGTGTAACACTATTGGTGCCAGAGATGCAATGCTGTGCTCATCTT 5014
Db 1694 CAATCTCTCTGGGACATCATGAATATTTCGTGAGTGTGGGCTGTGGTGTGGGCATCTT 1753
Qy 5015 TCTGGCTTCCAGCAGAGGCATATGTGGCCCTGCGCAACCTGCTCTCTCTCTCTCTCTT 5074
Db 1754 CATCGGTTTCAGAAAGAGCTACACTTCTCCAGAAACCTTCTCCCTCTGTGGCACT 1813
Qy 5075 GCTACTACTGTATGGCTGTGATGATCACACCGCTCATGTACCCAGCCCTCTTCTTCTTC 5134
Db 1814 GCTCCTGCTGTATGATGGCGGTCAITCCCATGATGTACCAGCATCTCTCTGTTTGA 1873
Qy 5135 GTGCCCCAGACAGCCTATGTGTGTCTCACTGATFAAACCTCTTTATTTGCAATCAATGG 5194
Db 1874 TGTCCCCAGCACAGCCTATGTGGCTTTTATCTGTGTCTTAATCTGTTCATCGGCATCAACAG 1933
Qy 5195 AAGCATGCCACCTTTGTGCTGAGCTTTT---CTCTGATCAGAACTGCGAGGAGGTGAG 5251
Db 1934 CAGTGCTATTACCTTCATCTTGGAAATTTATTTGAAATFAACCGGACGCTGTCTCAGGTTCAA 1993
Qy 5252 CCGGATCTTGAACAGAGTCTTCTTATCTTCTCCCACTTCTGTCTGGCGCGGGGCTCAT 5311
Db 1994 CGCGTCTGAGGAAGTGTCTCATTTGTTCTTCCCACTTCTGTCTGGCGCGGGGCTCAT 2053
Qy 5312 TGACATGGTGGGAACAGAGGCATGGCTGATGCTTTGAGGCTTTGGGAGACAGGCAATT 5371
Db 2054 TGACCTTGTGCACTGAGCAGGCTGTGACAGATGTCTATGCCCCGTTTGGTGGAGCACTC 2113
Qy 5372 CCAGTCAACCTCGCTGGGAGGTGTGCGCAAGAACCTCTTGGCCATGTTGATACAGGG 5431
Db 2114 TGCMAATCCGTTCCACTGGGACCTGATTTGGGAAGAACCTGTTTGGCAATGTTGGTGAAG 2173
Qy 5432 GCCCCTCTTCTTCTTCACTACTGCTGACAGCCAGCAACCTCTCTGCCACAGCC 5491
Db 2174 GGTGGTGTACTTCTCTGACCTTGTGGTCCAGCGCACTTCTTCTCTCCCAATGAT 2233
Qy 5492 CAGGCTGAGGTCTCTGCCATCTCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAACGGGA 5551
Db 2234 TGCCGAGCCCACTAAGGAGCCCATTTGTGATGAAGATGATGATGTGGCTGAAGAAAGACA 2293
Qy 5552 GCGGGTGGTCCAAAGAGCCACCCAGGGGATGTGTGTGCTGAGGAACCTTGACCAAGGT 5611

Db 2294 AAGATTTATTACTGGTGGAAATAAACTGACATCTTTAAGGCTACATGAACCAAGAT 2353
Qy 5612 ATACCGTGGGACAGAGATGCCAGCTGTTGACCGCTTGTGCTGGGGATTTCCCCCTGTGTA 5671
Db 2354 TTATCCAGGACCTCCAGCCAGCAGTGGACAGGCTGTGTGTCGAGTTCGCCCTGGAGA 2413
Qy 5672 GTGTTTGGGCTGCTGGTGTGAATGGAGCAGGGAAGACGCTCCACGTTTCCGATGGTAC 5731
Db 2414 GTCTTTGGCTCTCTGGAGTGTGTCGCGCAAAACCAACCAATTTCAAGATGCTCAC 2473
Qy 5732 GGGGACACATTTGCCAGCGGCGAGGCTGTGTCGCGAGGCCACAGCGTGGGCCCGGA 5791
Db 2474 TGGGACACACAGTGTACCTCAGGGGATGCCACCGTAGCAGGCAAGAGTATTTTAAACAA 2533
Qy 5792 ACCAGTGTGCGCACCTCAGCATGGGATAGTCCCTCAATCCGATGCCATCTTTAGCT 5851
Db 2534 TATTTCTGAAGTCCATCAAAATATGGCTACTGTCTCTCAGTTTGTATGCAATCGATGACT 2593
Qy 5852 GCTGACGGGCGGAGACCTGAGCTGTGTCGCGCTCGCGGCTGCGGCTGCCGAGGCCCA 5911
Db 2594 GCTCACAGGACGAAACATCTTTACCTTTATGCCCGCTTCGAGGTGTACAGCAGAAGA 2653
Qy 5912 GGTGTCACAGCCGCTGGCTCAGCGCTGGCGCTCTGGGACTCTCATGTGACGACGACCG 5971
Db 2654 AATCGAAAGGTTGCAAACTGGAGTATTAGAGCTTGGGCTGACTCTACGCCGACTG 2713
Qy 5972 GCCTGACGACCTTACAGCGGAGGGAACAAACGCAAGCTGGCGAGCGGCTTGGCGCTGGT 6031
Db 2714 CTTGGCTGGCACCTACAGTGGGGCAACAAAGCGGAAACTCTCCACAGCCATCGCACTCAT 2773
Qy 6032 TGGGACCCAGCCGCTGGTGTCTGGACGAGCCGACACAGGATGAGACCCCGCGCGG 6091
Db 2774 TGGTGGCCACCGCTGGTGTCTGGATGAGCCACACAGGAGTGAACCCCGAGGACG 2833
Qy 6092 GCGCTTCTTTGAAACAGCCCTTTTGGCCGTGTGGCGGAGGCGGCTTCAGTGATGCTCAC 6151
Db 2834 CCGCATGCTGTGNAACGTCACTGTGACATCATCAGAGAGGAGGAGGCTGTGGTCTCTCAC 2893
Qy 6152 CTCCCATAGCATGGAGAGTGTGAAGCGCTGTGCTGCGGCTAGCCATCATGTGTGAATGG 6211
Db 2894 ATCCACAGCATGGAAGATGTGAGGCACTGTGTACCCGCTGGCCATCATGTGTAAAGGG 2953
Qy 6212 GCGTTCGCTGCTGGGACCGCGCAACATCTCAAGGCGAGATTCGCGCGGCTCACAC 6271
Db 2954 CGCTTTTCGATGTATGGGACCAATTTCAGCATCTCAAGTCCAAATTTGAGATGCTTAT 3013
Qy 6272 ACTGACCTCGGGTGCCTCCCGCAAGTCCCGCGGAGCGGCTT-----CGTGGC 6325
Db 3014 CGTCACATGAAGATCAAAATCCCGAGAGGAGACCTGCTTCTTGAACCTGTGGA 3073
Qy 6326 GGCGGAGTTCCTGGTFCGAGCTGGCGGAGGACATGAGAGGCGGCTTGGCTTCCAGCT 6385
Db 3074 GCAGTTCCTCAGGGGAACTTCCAGGCACTGTGACAGGAGGAGGAGGCACTACAACTGCT 3133
Qy 6386 GCGG---CCGGGAGGCGCTGCGCCCTGCGCGCTTCTTGGAGAGCTGGCGGTGACGG 6442
Db 3134 CCAGTTCAGGTTCT 3193
Qy 6443 CGCAGAGCAGCGGCTGGAGGACTTTTTCGTCGAGCCAGCATGCTGGAGGAGGTATTCTT 6502
Db 3194 GGACGCTCTCATCGAGGAGTACTCAGTCACACAGACACACTGGACCGAGGTGTGT 3253
Qy 6503 GTACTTCTCAAGACCCAG 6521
Db 3254 AAATTTGTCTAAACAGCAG 3272

RESULT 8

AK052916

LOCUS

DEFINITION

AK052916 3605 bp mRNA linear HTC 03-APR-2004
Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched
library, clone:D830040G08 product:ATP-binding cassette, sub-family
A (ABCL), member 4, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK052916

AK052916.1 GI:26095451

HTC: CAP trapper.

Mus musculus (house mouse)

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

REFERENCE

5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3605)

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp).
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

FEATURES

source

URL: <http://fantom.gsc.riken.jp/>.

Location/Qualifiers

1..3605

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6J"

/db_xref="PANTOM_DB:D830040G08"

/db_xref="taxon:10090"

/clone="D830040G08"

/issue_type="heart"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days neonate"

1..3605

misc_feature

/note="ATP-binding cassette, sub-family A (ABC1), member 4

(MGD|GI:109424, GB|NM_007378, evidence: BLASTN, 99%,

match=3552)"

ORIGIN

Query Match 11.8%; Score 805.8; DB 3; Length 3605;
Best Local Similarity 55.8%; Pred. No. 7.2e-155;
Matches 1762; Conservative 0; Mismatches 1327; Indels 70; Gaps 9;

QY 3439 AGCTGCTGGCCCTGTACAGCACTGGGTGCGGGGACGGCTGTGTGAGGAGCTGCCAC 3498
DB 65 AGCTGATGGATTGGTATATACCACTATGCCAGAGGCAAGCTGGTGAATGCAITGGTC 124
QY 3499 ACGAGCTGCTGGTGTGCTGCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCT 3558
DB 125 AAGAACTTATCTTCTCTCTCCAAACAAGAATTTCAAGCAGAGACATATGCCAGCTTT 184
QY 3559 TCCGAGAGCTAGACACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3618
DB 185 TCCGAGAGCTGGAGGAGACTTGGCTGACTGGGCTCAGCAGCTTTGGAATTTCTGACA 244
QY 3619 CAGGCTCGAGAGATCTTCTGAAGGTGGTGGAGAGTGTCTCGGACACAGATATGG 3678
DB 245 CTCCCTGGAAGAGATTTTCTGAGGTGCACAGAGGATGCTGGAGCGGCTCTATGTTTG 304
QY 3679 AGGATGGCAGCTGGGGCAGCACCTATGCACAGGATTTGCTGGCTAGACGTAACCTGC 3738
DB 305 TAGTGGCGCTCAGCAGAAAGAGAACCGGCTCTCGGCACCTTGTTCGGCTCCCA 364
QY 3739 GGCTCAAGATCCGCCACAGGAGACAGCGCTGGAGAACGGGAAACAG-----CTGGGTC 3793
DB 365 CTGAGAGCTCAGGAGTATGCCAGGCCCCACATACCTGTTCCTCCAGGACAAAGTGATC 424
QY 3794 AGCCCCAGAGACTGACACAGGCTCTGGGCCACAGACCGCTGGGCCGGGTACAGGGTGGG- 3852
DB 425 CTCCCAAGGGCAACCTTCCCGAGAGCCAGAACCCCGGTGTCCCATTCACACAGGTG 484
QY 3853 ---CACTGACCGCCGAGCTCCAGGCGCTTCTCAAGGCTTTCTGTGGCCCTGGCC 3909
DB 485 CTGGCTGATTTCTCAACATGTGAGGCGCTGTAGTCAAGGATTCATCATCATCC 544
QY 3910 GCAGCGCGCGCGGCTGTTCGCCACAGATCGTGTGCTGCTGCTTCTGTGGCCCTGGCC 3969
DB 545 GCAGCGCAAGGACTTTGTGGCTCAGATTGTCTCCCTGCCACTTTCGTGTTTGGCTC 604
QY 3970 TCGTGTTCAGCTCATCTGTGCTCTCTTTGGGCACTACCCGGCTCTGGGCTCAGTCCCA 4029
DB 605 TGATGCTTCAATCATTTGTGATCCGTTGGTGAATTTCCAGCTTTGACTCTTCATCCCT 664
QY 4030 CCATGTACGGTGTCTCAGGTGCTTCTTCACTGAGGAGCGCCCGAGGACCTCGGACGTG 4089
DB 665 GGATGTATGGGATCAATATATACCTTCTTCAGATGGAGAACCCCAACATGAACACCTTG 724
QY 4090 CCCGCTCTCTCGAGCGCTGTGTGAGGAGGAGGAGCTGGAGAGC-----CCCCAGTGAG 4145
DB 725 AAGTACTGGCAGATGCTCTCTGTAACAGCCAGGCTTTGGCAACCGTTGTCTTAAGGAAG 784
QY 4146 CATAGTCTCCACAGGTTCTCGGCAACAGAGTGTCTGCTG-----4185
DB 785 AGTGGCTTCCGAGGTACCCATGATTAATGCAACCTCTCTGGAAGACGCCCTCCGTGTC 844

Qy	4186	---	AAGTGGCCAAAGGCTTTGGCCAGTGGCAACTGGACCCCAAGAGTCTCCATCCCAAGCCT	4242	Qy	5257	TCCTGAAACAGGCTTCTCTTATCTTCCCCCACTTCTGCTTGGGCGGGGGCTCATTTGACA	5316			
Db	845	CGAACATCACCCACCTGTTCAGAAAGCAGAAATGGACAGACGCCACCCCTCTCCCTCCT	904	Db	1925	TGTTGAGGAAAGTGTCTCATTTGTCTTCCCCCACTTCTGCTTGGGCGGGGGCTCATTTGACC	1984	Db	5317	TGGTGGCGAAACAGGCGCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGCAAGTTCAGT	5376
Qy	4243	GCAGTGTAGCCAGCCCGGTGCGCGCGCTGCTGCGCGGCTGCGCGGCTGCGAGCTGGTG	4302	Qy	5317	TGGTGGCGAAACAGGCGCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGCAAGTTCAGT	5376	Db	1985	TGGCATGAGCAAGCTGTGACAGAGCTCTATGCCAGTTTGGTGAGAGTACTCTGCAA	2044
Db	905	GCAAGTGAGCACCCAGAGAGAGCTGACCATGTTGCGCGAGTGGCCCGAGGGTCTGGAG	964	Db	1985	TGGCATGAGCAAGCTGTGACAGAGCTCTATGCCAGTTTGGTGAGAGTACTCTGCAA	2044	Qy	5377	CACCCCTCGCTGGAGGGTGGTGGCAAGAACCTTTGGGCAATGGTGATACAGGGGGCCCC	5436
Qy	4303	GTCCCCCTCGCCCCAGCAGTACCGGCTCTGGGGAAGTGGTTAGAAACCTGACAGGCC	4362	Qy	5377	CACCCCTCGCTGGAGGGTGGTGGCAAGAACCTTTGGGCAATGGTGATACAGGGGGCCCC	5436	Db	2045	ACCCATTCAGTGGGAGCTGATTTGGGAAGAACCTGGTGTGCTATGGCAATAGAGGGGTG	2104
Db	965	GGCTCCCAACCCACAGAGGACACAGCGCAGCACTGAAGTCTTACAGACCTCACGAACA	1024	Db	2045	ACCCATTCAGTGGGAGCTGATTTGGGAAGAACCTGGTGTGCTATGGCAATAGAGGGGTG	2104	Qy	5437	TCCTTCTTCTTTCACACTACTGTGTCAGCACCGCAAGCCAACTCTGCCACAGCCAGGG	5496
Qy	4363	GGAACTGTGTGACTTCTCTGTGTCAAGACCTTACCCGCGCTGGTGGCGGAGGCTGAAGA	4422	Qy	5437	TCCTTCTTCTTTCACACTACTGTGTCAGCACCGCAAGCCAACTCTGCCACAGCCAGGG	5496	Db	2105	TATATCTCTTCTGACCCCTCTCATCCAAACCAATTTTTTCTCACCCGGTGGATCGCTG	2164
Db	1025	GGAACTGTGTGACTTCTCTGTGTCAAGACCTTACCCGCGCTGGTGGCGGAGGCTGAAGA	1084	Db	2105	TATATCTCTTCTGACCCCTCTCATCCAAACCAATTTTTTCTCACCCGGTGGATCGCTG	2164	Qy	5497	TGAGGTCTCTGCCACTCTCGGAGAGGAGGACGAGGATGTAGCCCGTGAACGGGAGCGGG	5556
Qy	4423	CTAAGAAAGTGGTGAATGAGGTACAGGTACGAGGCTTCTCGCTGGGGGGCGAGACCCAG	4482	Qy	5497	TGAGGTCTCTGCCACTCTCGGAGAGGAGGACGAGGATGTAGCCCGTGAACGGGAGCGGG	5556	Db	2165	AGCTGTCTAGAGAGCCGTTTTTGTATGATGATGACGATGGCTGTAAGAAAGACAAAGAG	2224
Db	1085	GCAATTTCTGGGTCAATGAACAGAGGTATGAGGAATTTCCATCGGAGGAAATCTCCCTG	1144	Db	2165	AGCTGTCTAGAGAGCCGTTTTTGTATGATGATGACGATGGCTGTAAGAAAGACAAAGAG	2224	Qy	5557	TGFTCCAAGAGGACACCCAGGGGGATGTGTGTGTGTGAGGAACTTGACCAAGATATACC	5616
Qy	4483	GCCTGCC---CTCGGGCCAAAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGCGCTGCTGA	4539	Qy	5557	TGFTCCAAGAGGACACCCAGGGGGATGTGTGTGTGTGAGGAACTTGACCAAGATATACC	5616	Db	2225	TTATAGTGGGGGAAATAAAACCGACATCTTAAAGCTTAAACGAACCTAACCAAGTTTACT	2284
Db	1145	CTATCCCACTCAGCGGGGAGCACTGTGTGGCTTTCTAAGTGGCTTTGGCCAGATGATGA	1204	Db	2225	TTATAGTGGGGGAAATAAAACCGACATCTTAAAGCTTAAACGAACCTAACCAAGTTTACT	2284	Qy	5617	GTGGCAGAGGATGCCAGCTGTGTACCGCTTGTGCTGGGGATTCGCCCTGGTGGTGT	5676
Qy	4540	GTCCCTCGCTGGCGGGGCCCTCGACCGTGTCTGAAAAACCTCACAGCTGGGCTCACA	4599	Qy	5617	GTGGCAGAGGATGCCAGCTGTGTACCGCTTGTGCTGGGGATTCGCCCTGGTGGTGT	5676	Db	2285	CAGGCTCTTCAGTCCAGCAGTAGATCGGTATGTGTGTGGAGTTCGACCTCGGAGGTGCT	2344
Db	1205	ACGTGAGCGGGGTCTCTGTACACAGGAGGCGCTCCAAAGAAATGTTAGATTTCTCAAAC	1264	Db	2285	CAGGCTCTTCAGTCCAGCAGTAGATCGGTATGTGTGTGGAGTTCGACCTCGGAGGTGCT	2344	Qy	5677	TTGGGCTGTCTGGGTGTGAATGGAGCAGGGAAGAGCTCCACGTTTTGCGATGGTGCGGGG	5736
Qy	4600	GCCTGGAGTGTCAAGACAGTCTCAAGATCTGGTTCAACAACAAGAGGTGGCACTCCATGG	4659	Qy	5677	TTGGGCTGTCTGGGTGTGAATGGAGCAGGGAAGAGCTCCACGTTTTGCGATGGTGCGGGG	5736	Db	2345	TTGGCTCTCTGGGAGTGAATGGTGCAGGCAAAACCAACCATTCAGATGTCTACTGGG	2404
Db	1265	ATCTTGAACACACAGACAACTTAAGGTATGGTTTAAACAACAGGGCTGGCATGCCCTGG	1324	Db	2345	TTGGCTCTCTGGGAGTGAATGGTGCAGGCAAAACCAACCATTCAGATGTCTACTGGG	2404	Qy	5737	ACACATTTGGCCAGCAGGGGGCGAGGCTGTGTGTCAGAGCCACAGGCTGGCGGGAAACCCA	5796
Qy	4660	TGGCTTTTGTCAACCGAGCCAGCAACCAATCTCTCGTGTCTACCTGCCCCAGGCCCGG	4719	Qy	5737	ACACATTTGGCCAGCAGGGGGCGAGGCTGTGTGTCAGAGCCACAGGCTGGCGGGAAACCCA	5796	Db	2405	ACACCAAGTGCATCATCGGGGATGCTACTGTGGCAGGCAAGAGCATTTAACTAGTATTT	2464
Db	1325	TCAGCTTTCTGATGTGGTCTCAACAGCCACTTTTACGGGCCAGCTGCCAGGAGCAGGG	1384	Db	2405	ACACCAAGTGCATCATCGGGGATGCTACTGTGGCAGGCAAGAGCATTTAACTAGTATTT	2464	Qy	5797	GTGCTGGCAGCTCAGCATGGGATGCTCCCTCAATCCGATGCCATCTTTGAGCTGCTGA	5856
Qy	4720	CCGCGCAGCCACAGCATACACACTCAACCAACCCCTTGAACCTCACCCAGGAGCAGC	4779	Qy	5797	GTGCTGGCAGCTCAGCATGGGATGCTCCCTCAATCCGATGCCATCTTTGAGCTGCTGA	5856	Db	2465	CTGATGTCCATCAAAACATGGGCTACTGTCTCAGTTTGATGCAATCGATGACCTGCTCA	2524
Db	1385	ACCTTGAGGAGTATGGAATCACTGTCTATCGCCAGCCCTCTGAACCTCAGCAAGAGCAGC	1444	Db	2465	CTGATGTCCATCAAAACATGGGCTACTGTCTCAGTTTGATGCAATCGATGACCTGCTCA	2524	Qy	5857	CGGCGCCGAGCAGCTTGGAGCTGTCTGGCGCTTGGCGGCTGTCCCGAGGCCCAGGTTG	5916
Qy	4780	TGCTGAGGCTGCACTGATGGCTCTCTGGTGGAGCTCTCGTCTCCATCTGTGTGTCT	4839	Qy	5857	CGGCGCCGAGCAGCTTGGAGCTGTCTGGCGCTTGGCGGCTGTCCCGAGGCCCAGGTTG	5916	Db	2525	CAGGCAAGAACACCTTTTACTCTATGACAGGCTTGGCGGCGTGCATCGAAGGAATCG	2584
Db	1445	TCTGGACATCACAGTACTGACTACTTCCGTGGATGCGGTGGTGCATATGTGTGATTT	1504	Db	2525	CAGGCAAGAACACCTTTTACTCTATGACAGGCTTGGCGGCGTGCATCGAAGGAATCG	2584	Qy	5917	CCAGACCGCTGGCTCAGGCTTGGCGCTTGGAGCTCTCATGGTAGCAGACCGGCTG	5976
Qy	4840	TGCGCATGTCTTGTCCCGCCAGCTTCACTTTGTCTTCAATTAGAGAGAGTCAACC	4899	Qy	5917	CCAGACCGCTGGCTCAGGCTTGGCGCTTGGAGCTCTCATGGTAGCAGACCGGCTG	5976	Db	2585	AGAAGTTGCGCACTGGGGTATCCAGAGCCTGGGCTTGTCTCTACGCTGACCGCTGG	2644
Db	1505	TCGCGATGTCTCGTTCCAGCCAGCTTTGTCTTTACTTGTATCCAGGAGGGGTGACGA	1564	Db	2585	AGAAGTTGCGCACTGGGGTATCCAGAGCCTGGGCTTGTCTCTACGCTGACCGCTGG	2644	Qy	5977	CAGGCACTACAGCGGAGGAAACAAACGCAAGCTGGCGAGCGGCCCTGGCGTGTGGGG	6036
Qy	4900	GAGCAAGCACTGCAGCTCATGGGGGCGCTGTCCCGCCAGCTTACTTGGCTGGCAACT	4959	Qy	5977	CAGGCACTACAGCGGAGGAAACAAACGCAAGCTGGCGAGCGGCCCTGGCGTGTGGGG	6036	Db	2645	CAGGCACTACAGTGGAGGCAATAAGAGGAAACTCTCTACAGCCATAGCTCTCACTGGCT	2704
Db	1565	AAGCTAAACATTTGCAAGTTTATCAGTGGTGTGAGCGCCCAACCACTACTTGGCTGACCAACT	1624	Db	2645	CAGGCACTACAGTGGAGGCAATAAGAGGAAACTCTCTACAGCCATAGCTCTCACTGGCT	2704	Qy	6037	ACCAGCCGTGTGTGTTTTCTGAGCGAGCCGACACAGGATGGACCCAGCGCGCGGCT	6096
Qy	4960	TTCTCTGGGACATGTGTAACCTTACTTTGTGCGCAGCATGCTCGTGGTGTCTATCTTTCTGG	5019	Qy	6037	ACCAGCCGTGTGTGTTTTCTGAGCGAGCCGACACAGGATGGACCCAGCGCGCGGCT	6096	Db	2705	GCCCTCCCTGCTGCTGCTGATGAGCCACAGCAGGATGGATCCCGAGCAGCGCGCA	2764
Db	1625	TCCTCTGGGATTTATGAAATATGCTGTGAGTGGGATGGTGGTGGGCACTTCAATG	1684	Db	2705	GCCCTCCCTGCTGCTGCTGATGAGCCACAGCAGGATGGATCCCGAGCAGCGCGCA	2764	Qy	6097	TCCTTTGGAACAGCCTTTTGGCGTGTGCGGAGGCGGCTTCAAGTATGCTCACTCCC	6156
Qy	5020	CCTTCCAGCAGAGGGCATATGTGGCCCTCGCCAACTTCCCTGCTCTCTGCTGTGTGTAC	5079	Qy	6097	TCCTTTGGAACAGCCTTTTGGCGTGTGCGGAGGCGGCTTCAAGTATGCTCACTCCC	6156	Db	2765	TGCTGTGGAACACCATTTGTGAGCATCATCAGAAAGGAGAGCTGTGTGTTCTCACTCCC	2824
Db	1685	GATTTCAAAAGAAAGCCCTACACATCCCGACAAACCTTCCCTGCTCTGCTCTCTGCTCA	1744	Db	2765	TGCTGTGGAACACCATTTGTGAGCATCATCAGAAAGGAGAGCTGTGTGTTCTCACTCCC	2824	Qy	6157	ATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGGGGT	6216
Qy	5080	TACTGTATGGTGTGATCACAGCGTCTATGATACCGAGCTCTCTTCTTCTCCGTCG	5139	Qy	6157	ATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGGGGT	6216	Db	2825	ACAGCATGGAAGATGTAAGCTCTGTGTAGCGGCTGGCCATCATGTTGTAAGGGCACT	2884
Db	1745	TGCTGTATGGAATGGGAGTCAATTCCTATGATGATATCAGGATCTCTCTGTTGAAGTCC	1804	Db	2825	ACAGCATGGAAGATGTAAGCTCTGTGTAGCGGCTGGCCATCATGTTGTAAGGGCACT	2884	Qy	6217	TCCGCTGCTTGGGCGAGCC--CGCAACATCTCAAGGGCAGATTCGCGGGGTTCACACTG	6275
Qy	5140	CCAGCAGCCTATGTGTGTCTCACCTGCAATAAACCTCTTATTTGGCATCAATGGAAGCA	5199	Qy	6217	TCCGCTGCTTGGGCGAGCC--CGCAACATCTCAAGGGCAGATTCGCGGGGTTCACACTG	6275	Db	2885	TTCAAGTGTCTGGGTACCCATCCCAACCTCAAGTACAAAGTTTGGAGAGCGCTACATGTC	2944
Db	1805	CCAGCAGCCTATGTGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1864	Db	2885	TTCAAGTGTCTGGGTACCCATCCCAACCTCAAGTACAAAGTTTGGAGAGCGCTACATGTC	2944	Qy	6276	ACCTTGGGGTGGCGCGGCAAGGTCCAG-----CCGGCAGCGGCC	6317
Qy	5200	TGGCCACTTTTGTGCTTGAAGTCTTT---CTCTGATCAGAAAGCTCGAGGAGTGGCGGGA	5256	Qy	6276	ACCTTGGGGTGGCGCGGCAAGGTCCAG-----CCGGCAGCGGCC	6317	Db	2945	ACAATGAAATAAAATCTCAGAAAGGAGCACTTGTCTCCCGATTTGAATCTCTGTGAGAG	3004
Db	1865	CCATCAGCTTCTGCTTGAATTTATTGAGAATAACCGGACGCTGCTCAGGTTCAATGCCA	1924	Db	2945	ACAATGAAATAAAATCTCAGAAAGGAGCACTTGTCTCCCGATTTGAATCTCTGTGAGAG	3004	Qy	6318	TTCTGCGCGCGGAGTTTCCCTGCGAGCTGCGCGAGGACATGAGGAGCGGCTGCGC	6377

Accession	Sequence	Length
Db	TTCTTCCAGGGCAATTCCTTGGCAGGTGCAGAGGGAGAGACACCAAGCATGCTCCAG	3064
Qy	TTCCAGCTGCCTGGCGGAGGGCGCTGCCTGCCTGCCTGCGCTCTTTGAGAGCTGGCGGTG	6437
Db	TTCCAGTGCCTCG-----TCCTCCTGCGCAGGATCTTCCAGCTGCTCATTTCC	3115
Qy	CACGGCGCAGACGCGGTGGAGGACTTTTCCTGTAGCCAGACGATGCTGGAGGAGTTA	6497
Db	CACAAGACAGCTGCTCATCGAGGAGTACTCACTACCCAGACACCGCTGGACCAAGTG	3115
Qy	TTCTTGTACTTCTCCAGGACCAGGGGAAGCAGGAC	6536
Db	TTTGTAACTTCGCTAAACAGCAGACTGAGACCTTATGAC	3214

RESULT	9
BQ955522	
LOCUS	
DEFINITION	BQ955522 958 bp mRNA linear EST 21-AUG-2002
DEFINITION	AGENCOURT_8843758 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6456042 5' , mRNA sequence.

ACCESSION	BQ955522	
VERSION	BQ955522.1	GI:22371000
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 958)	
AUTHORS	NH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D.	

```

FEATURES
source
    1. .958
    Location/Qualifiers
        1. .958
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6456042"
        /tissue_type="epidermoid carcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_101"
        /notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match
Best Local Similarity 11.7%; Score 799; DB 5; Length 958;
Matches 892; Conservative 0; Mismatches 55; Indels 10; Gaps 5;

QY 2982 CAGGATGTGGGGCTGGTCTCCACGACGAGTGTGCAGATCGCCACCTCTCTGTGGTGGGATG 3041
|||||
Db 1 CAGGATGTGGGGCTGGTCTCCACGACGAGTGTGCAGATCGCCACCTCTCTGTGGTGGGATG 60
|||||

QY 3042 CAACGGAGCTGTCCTGTGGCCATTGCCCTTTGTGGGGCGGTCCCAAGTTGTTATCTCTGGAC 3101
|||||
Db 61 CAACGGAAGTGTCCGTGGGCATTGTGCTTTGTGGGGGGTCCCAAGTTGTTATCTCTGGAC 120
|||||

```

ORIGIN	NIN_MGC Library.									
Query Match	11.7%	Score 799;	DB 5;	Length 958;						
Best Local Similarity	93.2%	Pred. No. 1.4e-153;								
Matches 892;	Conservative 0;	Mismatches 55;	Indels 10;	Gaps 5;						
QY	2982	CAGGATGTGGGCTGGTCTCCAA	GACGAGTGTGCAGACTGCCACCTCTCTGGTGGGATG	3041						
Db	1	CAGGATGTGGGCTGGTCTCCAA	GACGAGTGTGCAGACTGCCACCTCTCTGGTGGGATG	60						
QY	3042	CAACGGAAGCTGTCCTGGGCCATGCTTTGTGGGCGGGTCCCAAGTTGTTATCTCTGGAC	3101							
Db	61	CAACGGAAGCTGTCCTGGGCCATGCTTTGTGGGCGGGTCCCAAGTTGTTATCTCTGGAC	120							

QY	3102	GAGCCTACGGCTGCGCTGGATCTCTGCTTCCGCGCGCGGTATTTGGGAGCTGCTGCTCAAA	3161
DB	121	GAGCCTACGGCTGCGCTGGATCTCTGCTTCCGCGCGCGGTATTTGGGAGCTGCTGCTCAAA	180
QY	3162	TACCGAGAAGGTCCACGCTGATCTCTCCACCCACACCTGGATGAGGCAGAGCTGCTG	3221
DB	181	TACCGAGAAGGTCCACGCTGATCTCTCCACCCACACCTGGATGAGGCAGAGCTGCTG	240
QY	3222	GGAGACCGTGTGGCTGTGGTGGCAGGTGGCCGCTTGTGTGCTGTGGCTCCCACTCTTC	3281
DB	241	GGAGACCGTGTGGCCGTGGTGGCAGGTGGCCGCTTGTGTGCTGTGGCTCCCACTCTTC	300
QY	3282	CTGGCCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCCGCCCTGCCCTG	3341
DB	301	CTGGCCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCCCGCCCTGCCCTG	360
QY	3342	ACCACCAATGAGAAGGCTGACACTGACATGGAGGGCAGTGTGGACACCAGCAGGAAAAG	3401
DB	361	ACCACCAATGAGAAGGCTGACACTGACATGGAGGGCAGTGTGGACACCAGCAGGAAAAG	420
QY	3402	AAGAATGGCAGCCAGGGCAGAGAGTCGGCACTCTCAGCTGTGGCCCTGGTACAGCAC	3461
DB	421	AAGAATGGCAGCCAGGGCAGAGAGTCGGCACTCTCAGCTGTGGCCCTGGTACAGCAC	480
QY	3462	TGGGTGCCCGGGCAGCGCTGGTGGAGGAGCTGCCACACGAGCTGTGTGCTGGTGGCCCC	3521
DB	481	TGGGTGCCCGGGCAGCGCTGGTGGAGGAGCTGCCACACGAGCTGTGTGCTGGTGGCCCC	540
QY	3522	TACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCGAGAGCTAGACACGCGGCTG	3581
DB	541	TACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCGAGAGCTAGACACGCGGCTG	600
QY	3582	GCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCCTCGAGGAGATCTTCTCTG	3641
DB	601	GCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCCTCGAGGAGATCTTCTCTG	660
QY	3642	AAGTGTGTGAGGAGTGTGTGCGGACACAGATATGAGAGATGCGAGCTCGGGCAGCAC	3701
DB	661	AAGTGTGTGAGAGATGTGTGCGGACCCAGATATGAGAGATGCGAGCTCGGGCAGCAC	720
QY	3702	CTATGACACAGGCA--TTGCTGGCTTAGACGTAAACCTTCGCGCTCAAGATCCGCCACAGGA	3760
DB	721	CTATGACACAGCATTTGTGGCTTAAACCGTAAACCGTACCCCTACCGCTCCAGATGCGGGCAGGA	780
QY	3761	GACAGCGCTGG--AGAAACGGGGAACAGCTGGGTGAGCCCCAGAGACTGACCAAGGG--CT	3816
DB	781	GACAGCGCTGGGAGAACGGGGAACCAACTTGTGTGAGCCCCAGAGACTGACCAAGGGGCTC	840
QY	3817	CTGGGCCAGACGCGCTGGGCGGGGTACA--GGGCTTGGGCACCTGACCCCGCAGACAGCTCC	3873
DB	841	TGGGGCCAAACCCCGGGGGGCCGGTTAAACGGGGCTGGGAGCCGAACCCGCCACCACTCC	900
QY	3874	AGGCCCTGTCTTCAAGCGCTTTT--CTGCTTGCCCGCCGACGCCCGCGGCGCTGTTT	3928
DB	901	AGGCCCTGTGTTTAAAGGCTTTTCTGTGCTTGCCCCCCCCAAACCGCGGGGCGGT	957

RESULT 10	BC036566	1934 bp	mRNA	linear	HTC 04-MAR-2003
LOCUS	Mus musculus, clone IMAGE:1366813,				
DEFINITION	BC036566				
ACCESSION	BC036566				
VERSION	BC036566.1	GI:23306463			
KEYWORDS	HTC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1934)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-AUG-2002) National Institutes of Health, Mammalian				

IMAGE:6537923 5', mRNA sequence.
BU528031 GI:22838472
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 914)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2701 row: k column: 11
High quality sequence stop: 602.
Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6537923"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Query Match 11.0%; Score 751.8; DB 5; Length 914;
Best Local Similarity 94.5%; Pred. No. 7.5e-144;
Matches 834; Conservative 0; Mismatches 42; Indels 7; Gaps 5;
QY 2982 CAGGATGGGGCTGCTCCAGCAGAGTGTGCAGACTCGCCACCTCTCTGGTGGGATG 3041
DB 1 CAGGATGGGGCTGCTCCAGCAGAGTGTGCAGACTCGCCACCTCTCTGGTGGGATG 60
QY 3042 CAACGGAAGCTGTCCGTCGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTCGAC 3101
DB 61 CAACGGAAGCTGTCCGTCGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTCGAC 120
QY 3102 GAGCTACGGTGGCGTGGATCTGCTTCCCGCCCGGTATTTGGAGCTGCTGCTCAA 3161
DB 121 GAGCTACGGTGGCGTGGATCTGCTTCCCGCCCGGTATTTGGAGCTGCTGCTCAA 180
QY 3162 TACCGAGAGTGTGCAGCTGATCTCTCCACCCACCTGGATGAGGAGAGTGTGCTG 3221
DB 181 TACCGAGAGTGTGCAGCTGATCTCTCCACCCACCTGGATGAGGAGAGTGTGCTG 240
QY 3222 GGAGACCTGTGGCTGTGGCGAGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 3281
DB 241 GGAGACCTGTGGCGAGTGGCGAGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 300
QY 3282 CTGCGCCGTACCTGGGCTCCGGCTACTACTGAGCTGTGTGAAGGCCCGGCTGCCCGCTG 3341
DB 301 CTGCGCCGTACCTGGGCTCCGGCTACTACTGAGCTGTGTGAAGGCCCGGCTGCCCGCTG 360
QY 3342 ACCACCAATGAGAGGCTGACACTGACATGAGGGGAGTGTGGACACAGGAGGAGAAAG 3401
DB 361 ACCACCAATGAGAGGCTGACACTGACATGAGGGGAGTGTGGACACAGGAGGAGAAAG 420

QY 3402 AAGATGCGAGCCAGGGCAGCAGAGTGTGGCGCTCTCTCAGCTGTGSCCTGTGTACAGCAC 3461
DB 421 AAGATGCGAGCCAGGGCAGCAGAGTGTGGCGCTCTCTCAGCTGTGSCCTGTGTACAGCAC 480
QY 3462 TGGGTGCCCCGGGCGACGGCTGTGGTGGAGAGCTGCCACAGAGCTGTGTGTGTGTGCTG 3521
DB 481 TGGGTGCCCCGGGCGACGGCTGTGGTGGAGAGCTGCCACAGAGCTGTGTGTGTGTGCTG 540
QY 3522 TACACGGGTGCCCCATGACGGCAGCTTCCGCACACTTTCGAGAGCTAGACACGGGCTG 3581
DB 541 TACACGGGTGCCCCATGACGGCAGCTTCCGCACACTTTCGAGAGCTAGACACGGGCTG 600
QY 3582 CGGAGCTGTAGGCTCACTGGCTACGGGATCTCCGACACAGCTCGAGAGATCTT-CCT 3640
DB 601 CGGAGCTGTAGGCTCACTGGCTACGGGATCTCCGACACAGCTCGAGAGATCTTCCCT 660
QY 3641 GAAAGTGTGGAGAGTGTGTGGCGACACAGATATGGAGATGCGAGCTGCGGGCAGCA 3700
DB 661 GAAAGTGTGGAGAGTGTGTGGCGACACAGATATGGAGATGCGAGCTGCGGGCAGCA 720
QY 3701 CCTATGACAGGCTTGTGGCTAGAGT-AACCTGGGCTCAAGATGCGGCAC-AG 3758
DB 721 CCTATGACAGGCTTGTGGCTAGAGTAAACCTAGAGTAAACCTAGAGTGGCGCCCAAG 780
QY 3759 GAGACAGCG-CTGGAGAACGGGGAACC---AGCTGGGTGAGCCCGCAGAGACTGACAGGG 3814
DB 781 GAGACAGCGCTGTGAGAACGGGGAACCAGCTGGTGGCGCCCGCAGAGACTGACAGGG 840
QY 3815 CTCTGGGCGACAGCCCGTGGCGCGGTACAGGGTGGGCACTG 3857
DB 841 CTCTGGGCGCGAGACCCCGTGGGCGCGGGAACAGGCGCTTG 883

RESULT 12
BU846021
LOCUS
DEFINITION AGNCOURT 10413059 NIH MGC 109 Homo sapiens cDNA clone
IMAGE:6579963 5', mRNA sequence.
ACCESSION BU846021
VERSION BU846021.1 GI:24030726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 969)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2784 row: c column: 03
High quality sequence start: 23
High quality sequence stop: 473.
Location/Qualifiers
1..969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6579963"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned

FEATURES
source

Db 554 CGCCTAGCATCATGGTGAATGGCGGTTCCGCTCGCTGGGAGCCCGCAACATCTCAAG 495
Qy 6249 GGCAGATTCCGCGCGGTGACACACTGA-CCCTGGGGTGGCGCGCAAGTCCAGCC 6307
Db 494 GGCAGATTCCGCGCGGTGACACACTGA-CCCTGGGGTGGCGCGCAAGTCCAGCC 435
Qy 6308 GGCAGCGGCTTTCGTGGCGCGAGTTCCTGGGTGGAGTGGCGAGGACATGGAGG 6367
Db 434 GGCAGCGGCTTTCGTGGCGCGAGTTCCTGGGTGGAGTGGCGAGGACATGGAGG 375
Qy 6368 CGCCTCGGCTTCAGCTGCGCGGAGGCGCTGGCGCCCTGGCGCGGCTTTCGGAGA 6427
Db 374 CGCCTCGGCTTCAGCTGCGCGGAGGCGCTGGCGCCCTGGCGCGGCTTTCGGAGA 315
Qy 6428 GCTGGCGGTGACGCGGAGAGACGCGGTGGAGACTTTTCGTGGAGCAGACAGATGCT 6487
Db 314 GCTGGCGGTGACGCGGAGAGACGCGGTGGAGACTTTTCGTGGAGCAGACAGATGCT 255
Qy 6488 GGAGGAGGTATTTCTTGTACTTCTCAAGGACGAGGAGGAGGACACCGAAGAGCA 6547
Db 254 GGAGGAGGTATTTCTTGTACTTCTCAAGGACGAGGAGGAGGACACCGAAGAGCA 195
Qy 6548 GAAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 6607
Db 194 GAAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 135
Qy 6608 CCAGTTCCTCGATGACCTTAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 6667
Db 134 CCAGTTCCTCGATGACCTTAGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 75
Qy 6668 GGCCTGGGAGGCGCTTGGGAAATGGCAAGGCAAGGTAGAGTGGCTTAGAGGCGCTGGACT 6727
Db 74 GGCCTGGGAGGCGCTTGGGAAATGGCAAGGCAAGGTAGAGTGGCTTAGAGGCGCTGGAGCC--TGACT 17
Qy 6728 CAGGCTGGCAGAGGGG 6743
Db 16 CAGGCTGGCAGAGGGG 1

RESULT 14
CD629699/c
LOCUS 55052443J1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD629699
ACCESSION CD629699
VERSION CD629699.1 GI:40277965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 749)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J., and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1..749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 10.5%; Score 717.4; DB 6; Length 749;
Best Local Similarity 98.9%; Pred. No. 8.7e-137;
Matches 733; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

Qy 4969 ACATGTGTAACACTTCTGGTGCCA--GCATGCACTGCTGGTGCTCATCTTTCTGGCTTCCA 5026
Db 741 ACATGTGTAACACTTCTGGTGCCAAGCATCTCTGGTGCTCATCTTTCTGGCTTCCA 682
Qy 5027 GCAGAGGAGATATGTGGCCCTTGCACACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5086
Db 681 GCAGAGGAGATATGTGGCCCTTGCACACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 622
Qy 5087 TGGCTGCTGATACACAGCTCATGTACCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5146
Db 621 TGGCTGCTGATACACAGCTCATGTACCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 562
Qy 5147 AGCTATGTGCTGCTCACTGTCATAAACCTCTTTATGTCATCAATGGAAGCATGGCCAC 5206
Db 561 AGCTATGTGCTGCTCACTGTCATAAACCTCTTTATGTCATCAATGGAAGCATGGCCAC 502
Qy 5207 TTTTGTGCTGAGCTCTTCTCTGATCAAGAGCTGAGGAGTGAGCCGGATCTTGAACA 5266
Db 501 TTTTGTGCTGAGCTCTTCTCTGATCAAGAGCTGAGGAGTGAGCCGGATCTTGAACA 442
Qy 5267 GGTCTTCTTATCTTCCCTTCT 5326
Db 441 GGTCTTCTTATCTTCCCTTCT 382
Qy 5327 CCAGGCTATGCTGATGCTTTTGGAGCTTTGGGAGACAGGCAAGTTCAGTCAACCTCTGG 5386
Db 381 CCAGGCTATGCTGATGCTTTTGGAGCTTTGGGAGACAGGCAAGTTCAGTCAACCTCTGG 322
Qy 5387 TTGGAGGTGCTGCGCAAGAACCTCTTGGCCATGATGATACAGGCGCCCTCTCTCTCTCT 5446
Db 321 TTGGAGGTGCTGCGCAAGAACCTCTTGGCCATGATGATACAGGCGCCCTCTCTCTCTCT 262
Qy 5447 CTTCACTACTGCTGCGACCGAGCAACTCTGCGCACAGCCAGGCTGAGTCTCT 5506
Db 261 CTTCACTACTGCTGCGACCGAGCAACTCTGCGCACAGCCAGGCTGAGTCTCT 202
Qy 5507 GCCACTCTGGGAGGAGGAGGAGATGTAGCCCTGAAACGGGAGCGGCTGCTCAAGG 5566
Db 201 GCCACTCTGGGAGGAGGAGGAGGAGATGTAGCCCTGAAACGGGAGCGGCTGCTCAAGG 142
Qy 5567 AGCCACCCAGGCGGATGTTGGTGTCTGAGGAACTTGACCAAGGTATACCGTGGGAGAG 5626
Db 141 AGCCACCCAGGCGGATGTTGGTGTCTGAGGAACTTGACCAAGGTATACCGTGGGAGAG 82
Qy 5627 GATGCCAGCTCTGACCGCTTGTGCTGGGAGTTCCTGCTGAGTCTTTGGGCTGCT 5686
Db 81 GATGCCAGCTCTGACCGCTTGTGCTGGGAGTTCCTGCTGAGTCTTTGGGCTGCT 22
Qy 5687 GGGTGTGAATGGAGCAGGGAA 5707
Db 21 GGGTGTGAATGGAGCAGGGAA 1

RESULT 15
CD629709
LOCUS 55058387J1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD629709
ACCESSION CD629709
VERSION CD629709.1 GI:40277975
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 730)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J., and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102

FEATURES Email: gfu@incyte.com.
 Location/Qualifiers
 1. .730
source /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="FLP"
 /note="Vector: pDrive Cloning Vector"

ORIGIN

```
Query Match      10.4%; Score 704.8; DB 6; Length 730;
Best Local Similarity 99.5%; Pred. No. 3.4e-134;
Matches 728; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 3395 GGAAGAAGAAATGGCAGCCAGGACGAGCTCGGCACCTCTCAGCTGTGGCCCTGGT 3454
Db 1 GGAAGAAGAAATGGCAGCCA-GGCAGCAGAGTCGGCACTCCTCAGCTGTGGCCCTGGT 59

Qy 3455 ACAGCACTGGTGCCTGGGCGACGGCTGGTGGAGAGCTGCCACACGAGCTGGTGGT 3514
Db 60 ACAGCACTGGTGCCTGGGCGACGGCTGGTGGAGAGCTGCCACACGAGCTGGTGGT 119

Qy 3515 GCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTAGACAC 3574
Db 120 GCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTAGACAC 179

Qy 3575 GCGGCTGGGAGCTGAGGCTCACTGGCTACGGGATCTCGACACCAAGCTTCGAGGAGAT 3634
Db 180 GCGGCTGGGAGCTGAGGCTCACTGGCTACGGGATCTCGACACCAAGCTTCGAGGAGAT 239

Qy 3635 CTTCCTGAAGTGTGGAGAGTGTCTGGGACACAGATATGGAGATGGCAGCTGGG 3694
Db 240 CTTCCTGAAGTGTGGAGAGTGTCTGGGACACAGATATGGAGATGGCAGCTGGG 299

Qy 3695 GCAGCACTATGCACAGGCTTGTGGCTAGAGCTAACCTGGCGCTCAAGATGCCGCC 3754
Db 300 GCAGCACTATGCACAGGCTTGTGGCTAGAGCTAACCTCAAGATGCCGCC 359

Qy 3755 ACAGGACACGCTGGAGAACCGGGAAACAGCTGGGTACGCCCCAGAGACTGACAGGG 3814
Db 360 ACAGGACACGCTGGAGAACCGGGAAACAGCTGGGTACGCCCCAGAGACTGACCA-GG 418

Qy 3815 CTCTGGGCCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGACAGAGCTCCA 3874
Db 419 CTCTGGGCCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGAGCCGCGAGCTCCA 478

Qy 3875 GGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCGCGCAGCGCGCGGCTGTTCGCCCA 3934
Db 479 GGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCGCGCAGCGCGCGGCTGTTCGCCCA 538

Qy 3935 GATCGTGTGCTGCTGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 3994
Db 539 GATCGTGTGCTGCTGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 598

Qy 3995 TTTCCGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTTGCTT 4054
Db 599 TTTCCGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGGTGTCTCAGGTTGCTT 658

Qy 4055 CTTCAGTGAAGACGCCCGGAGCCCTGGACGTGCCCGGCTGCTCGAGGCGCTGCTGCA 4114
Db 659 CTTCAGTGAAGACGCCCGGAGCCCTGGACGTGCCCGGCTGCTCGAGGCGCTGCTGCA 718

Qy 4115 GGAGGCAGGACT 4126
Db 719 GGAGGCAGGACT 730
```

Search completed: December 30, 2004, 05:09:45
Job time : 18920 secs